



101	19	95.0	241	4	US-09-134-001C-5598	Sequence 5598, Ap	174	19	95.0	337	4	US-09-371-671B-2	Sequence 2, Appli
102	19	95.0	241	4	US-09-489-039A-7795	Sequence 7795, Ap	175	19	95.0	338	4	US-09-328-352-7151	Sequence 7151, Ap
103	19	95.0	243	4	US-09-134-001C-3587	Sequence 3587, Ap	176	19	95.0	339	3	US-09-345-468-3	Sequence 3, Appli
104	19	95.0	246	4	US-09-252-991A-30976	Sequence 30976, A	177	19	95.0	339	4	US-09-414-453A-3	Sequence 3, Appli
105	19	95.0	247	4	US-09-230-196-16	Sequence 16, Appl	178	19	95.0	339	4	US-09-134-000C-5209	Sequence 5209, Ap
106	19	95.0	247	4	US-09-372-422A-48	Sequence 48, Appl	179	19	95.0	342	4	US-09-543-681A-4207	Sequence 4207, Ap
107	19	95.0	249	3	US-09-154-083-2	Sequence 2, Appli	180	19	95.0	345	4	US-09-107-532A-3849	Sequence 3849, Ap
108	19	95.0	249	3	US-09-345-468-9	Sequence 9, Appli	181	19	95.0	346	4	US-09-260-965-107	Sequence 107, App
109	19	95.0	249	3	US-09-372-422A-22	Sequence 22, Appl	182	19	95.0	347	4	US-09-134-001C-3298	Sequence 3298, Ap
110	19	95.0	249	4	US-09-414-453A-9	Sequence 9, Appli	183	19	95.0	350	4	US-09-655-270A-17	Sequence 17, Appl
111	19	95.0	249	4	US-09-352-991A-22610	Sequence 22610, A	184	19	95.0	350	4	US-09-651-941-21	Sequence 21, Appl
112	19	95.0	250	1	US-08-234-939-2	Sequence 2, Appli	185	19	95.0	350	4	US-09-955-597-21	Sequence 21, Appl
113	19	95.0	250	1	US-08-558-865-2	Sequence 2, Appli	186	19	95.0	353	4	US-09-253-991A-37528	Sequence 27528, A
114	19	95.0	250	3	US-08-654-025-7	Sequence 7, Appli	187	19	95.0	356	4	US-09-328-352-7671	Sequence 7671, Ap
115	19	95.0	250	3	US-08-654-025-7	Sequence 7, Appli	188	19	95.0	357	4	US-09-540-236-3808	Sequence 3808, Ap
116	19	95.0	251	4	US-09-252-991A-23816	Sequence 23816, A	189	19	95.0	363	4	US-09-252-991A-30015	Sequence 30015, A
117	19	95.0	253	4	US-09-252-991A-17352	Sequence 17352, A	190	19	95.0	364	4	US-09-543-681A-7525	Sequence 7525, Ap
118	19	95.0	253	4	US-09-107-532A-5072	Sequence 5072, Ap	191	19	95.0	368	4	US-09-252-991A-24621	Sequence 24621, Ap
119	19	95.0	254	4	US-09-372-422A-34	Sequence 34, Appl	192	19	95.0	368	4	US-09-489-039A-9556	Sequence 9556, Ap
120	19	95.0	257	2	US-08-506-240A-3	Sequence 3, Appli	193	19	95.0	370	4	US-09-540-224-4	Sequence 4, Appli
121	19	95.0	257	4	US-09-543-681A-6811	Sequence 6811, Ap	194	19	95.0	370	4	US-09-564-595D-53	Sequence 53, Appli
122	19	95.0	259	4	US-09-543-681A-6390	Sequence 6390, Ap	195	19	95.0	370	4	US-09-808-972-4	Sequence 4, Appli
123	19	95.0	259	4	US-09-489-039A-11936	Sequence 11936, A	196	19	95.0	371	2	US-08-837-593-6	Sequence 6, Appli
124	19	95.0	261	4	US-09-352-991A-23795	Sequence 23795, A	197	19	95.0	372	4	US-09-800-729-213	Sequence 213, App
125	19	95.0	262	4	US-09-352-991A-19738	Sequence 19738, A	198	19	95.0	374	4	US-09-489-039A-12678	Sequence 12678, A
126	19	95.0	263	4	US-09-800-729-88	Sequence 88, Appl	199	19	95.0	388	4	US-09-489-039A-8170	Sequence 8170, Ap
127	19	95.0	269	1	US-08-447-554-5	Sequence 5, Appli	200	19	95.0	390	4	US-09-933-313B-8	Sequence 8, Appli
128	19	95.0	269	1	US-08-468-763-17	Sequence 17, Appl	201	19	95.0	390	4	US-08-311-731A-332	Sequence 332, App
129	19	95.0	269	1	US-08-468-160-5	Sequence 5, Appli	202	19	95.0	390	4	US-09-543-681A-6415	Sequence 6415, Ap
130	19	95.0	269	2	US-09-393-996A-17	Sequence 17, Appl	203	19	95.0	390	4	US-09-489-039A-13547	Sequence 13547, A
131	19	95.0	270	4	US-09-252-991A-21912	Sequence 21912, A	204	19	95.0	391	2	US-08-928-692-26	Sequence 26, Appl
132	19	95.0	271	4	US-09-107-532A-5071	Sequence 5071, Ap	205	19	95.0	391	4	US-09-339-972-26	Sequence 26, Appl
133	19	95.0	273	4	US-09-489-039A-9190	Sequence 9190, Ap	206	19	95.0	394	4	US-09-934-903-4	Sequence 4, Appli
134	19	95.0	277	1	US-08-400-413-1	Sequence 1, Appli	207	19	95.0	397	4	US-09-252-991A-22235	Sequence 22235, A
135	19	95.0	280	1	US-08-595-559-3	Sequence 3, Appli	208	19	95.0	398	4	US-09-242-859A-4	Sequence 4, Appli
136	19	95.0	281	4	US-09-314-701-58	Sequence 58, Appl	209	19	95.0	398	4	US-09-242-859A-8	Sequence 8, Appli
137	19	95.0	282	4	US-09-134-000C-3572	Sequence 3572, Ap	210	19	95.0	398	4	US-09-252-991A-17379	Sequence 17379, A
138	19	95.0	283	4	US-08-956-171B-5203	Sequence 5203, Ap	211	19	95.0	400	2	US-08-733-825-2	Sequence 2, Appli
139	19	95.0	288	4	US-09-252-991A-18721	Sequence 18721, A	212	19	95.0	400	3	US-09-264-097-6	Sequence 6, Appli
140	19	95.0	290	4	US-09-743-847-2	Sequence 2, Appli	213	19	95.0	403	4	US-09-540-236-2573	Sequence 2573, Ap
141	19	95.0	291	4	US-09-252-991A-25517	Sequence 25517, A	214	19	95.0	405	3	US-09-232-200-63	Sequence 63, Appl
142	19	95.0	291	4	US-09-489-039A-9578	Sequence 9578, Ap	215	19	95.0	405	4	US-09-232-197-63	Sequence 63, Appl
143	19	95.0	294	4	US-09-523-263B-20	Sequence 20, Appl	216	19	95.0	405	4	US-09-291-023A-20	Sequence 20, Appl
144	19	95.0	295	4	US-09-543-681A-5560	Sequence 5560, Ap	217	19	95.0	405	4	US-09-232-201-63	Sequence 63, Appl
145	19	95.0	297	4	US-09-489-039A-9087	Sequence 9087, Ap	218	19	95.0	405	4	US-09-540-715A-20	Sequence 20, Appl
146	19	95.0	299	4	US-09-314-701-46	Sequence 46, Appl	219	19	95.0	405	4	US-09-232-195-63	Sequence 63, Appl
147	19	95.0	299	4	US-09-393-634-35	Sequence 35, Appl	220	19	95.0	406	4	US-09-543-681A-7962	Sequence 7962, Ap
148	19	95.0	301	4	US-09-489-039A-11964	Sequence 11964, A	221	19	95.0	410	2	US-08-723-415B-10	Sequence 10, Appl
149	19	95.0	311	4	US-09-614-912-198	Sequence 198, App	222	19	95.0	410	2	US-08-723-415B-11	Sequence 11, Appl
150	19	95.0	317	4	US-09-485-077A-17	Sequence 17, Appl	223	19	95.0	410	2	US-08-428-131-2	Sequence 2, Appli
151	19	95.0	318	2	US-08-872-719-2	Sequence 2, Appli	224	19	95.0	410	2	US-08-602-846-2	Sequence 2, Appli
152	19	95.0	318	3	US-08-957-302A-12	Sequence 12, Appl	225	19	95.0	410	3	US-09-078-596-2	Sequence 2, Appli
153	19	95.0	318	3	US-09-336-890-2	Sequence 2, Appl	226	19	95.0	410	3	US-09-189-627A-10	Sequence 10, Appl
154	19	95.0	318	3	US-09-542-403-12	Sequence 12, Appl	227	19	95.0	410	3	US-09-189-627A-11	Sequence 11, Appl
155	19	95.0	318	4	US-09-568-499-2	Sequence 2, Appli	228	19	95.0	410	4	US-09-710-861-10	Sequence 10, Appl
156	19	95.0	318	4	US-09-976-594-938	Sequence 938, App	229	19	95.0	410	4	US-09-710-861-11	Sequence 11, Appl
157	19	95.0	319	2	US-08-795-927-4	Sequence 4, Appli	230	19	95.0	410	4	US-09-252-991A-24839	Sequence 24839, A
158	19	95.0	319	3	US-09-345-468-5	Sequence 5, Appli	231	19	95.0	411	4	US-09-540-236-3549	Sequence 3549, Ap
159	19	95.0	319	4	US-09-414-453A-5	Sequence 5, Appli	232	19	95.0	414	4	US-09-252-991A-17176	Sequence 17176, A
160	19	95.0	319	4	US-09-242-991A-18807	Sequence 18807, A	233	19	95.0	415	4	US-09-134-001C-5101	Sequence 5101, Ap
161	19	95.0	321	4	US-09-489-039A-9512	Sequence 9512, Ap	234	19	95.0	415	4	US-09-252-991A-31684	Sequence 31684, A
162	19	95.0	323	4	US-09-232-991A-17111	Sequence 17111, A	235	19	95.0	415	4	US-09-252-991A-26713	Sequence 26713, A
163	19	95.0	323	4	US-09-543-681A-6957	Sequence 6957, Ap	236	19	95.0	424	4	US-09-328-352-4187	Sequence 4187, Ap
164	19	95.0	323	4	US-09-489-039A-7408	Sequence 7408, Ap	237	19	95.0	424	4	US-09-489-039A-9628	Sequence 9628, Ap
165	19	95.0	324	4	US-09-338-352-4636	Sequence 4636, Ap	238	19	95.0	425	4	US-09-489-039A-8386	Sequence 8386, Ap
166	19	95.0	326	4	US-09-230-196-20	Sequence 20, Appl	239	19	95.0	431	4	US-09-134-001C-4257	Sequence 4257, Ap
167	19	95.0	326	4	US-09-489-039A-9512	Sequence 9512, Ap	240	19	95.0	431	4	US-09-252-991A-24878	Sequence 24878, A
168	19	95.0	327	4	US-09-252-991A-30065	Sequence 30065, A	241	19	95.0	432	1	US-08-522-166-8	Sequence 8, Appli
169	19	95.0	328	4	US-09-489-039A-9306	Sequence 9306, Ap	242	19	95.0	432	2	US-08-488-382A-8	Sequence 8, Appli
170	19	95.0	331	4	US-09-489-039A-9639	Sequence 9639, Ap	243	19	95.0	432	2	US-08-488-382A-8	Sequence 8, Appli
171	19	95.0	335	1	US-08-202-054-2	Sequence 2, Appli	244	19	95.0	435	4	US-09-252-991A-30562	Sequence 30562, A
172	19	95.0	335	1	US-08-446-923-2	Sequence 2, Appli	245	19	95.0	436	4	US-09-161-994A-9	Sequence 9, Appli
173	19	95.0	335	4	US-09-482-973-118	Sequence 118, App	246	19	95.0	436	4	US-09-134-000C-3950	Sequence 3950, Ap



247	19	95.0	436	6	5405943-4	Patent No. 5405943	320	19	95.0	485	4	US-09-540-715A-18	Sequence 18, Appl
248	19	95.0	437	4	US-09-252-991A-25332	Sequence 25332, A	321	19	95.0	485	4	US-09-540-715A-19	Sequence 19, Appl
249	19	95.0	438	4	US-09-252-991A-22487	Sequence 22487, A	322	19	95.0	485	4	US-09-769-864-1	Sequence 1, Appl
250	19	95.0	444	4	US-09-252-991A-19256	Sequence 19256, A	323	19	95.0	485	4	US-09-769-864-2	Sequence 2, Appl
251	19	95.0	444	4	US-09-252-991A-19256	Sequence 19256, A	324	19	95.0	485	4	US-09-769-864-6	Sequence 6, Appl
252	19	95.0	444	4	US-09-252-991A-25986	Sequence 25986, A	325	19	95.0	485	4	US-09-769-864-7	Sequence 7, Appl
253	19	95.0	445	4	US-09-107-532A-7232	Sequence 7232, Ap	326	19	95.0	486	4	US-09-769-864-8	Sequence 8, Appl
254	19	95.0	445	4	US-09-328-352-6076	Sequence 6076, Ap	327	19	95.0	486	4	US-09-381-687-4	Sequence 4, Appl
255	19	95.0	445	4	US-09-134-000C-5444	Sequence 5444, Ap	328	19	95.0	493	3	US-08-996-441B-70	Sequence 70, Appl
256	19	95.0	450	4	US-09-543-681A-7260	Sequence 7260, Ap	329	19	95.0	493	3	US-08-993-722A-70	Sequence 70, Appl
257	19	95.0	451	4	US-09-134-000C-6421	Sequence 6421, Ap	330	19	95.0	493	3	US-08-993-170A-70	Sequence 70, Appl
258	19	95.0	455	4	US-09-252-991A-19356	Sequence 19356, A	331	19	95.0	493	3	US-08-993-775B-70	Sequence 70, Appl
259	19	95.0	456	4	US-09-252-991A-31884	Sequence 31884, A	332	19	95.0	493	3	US-09-177-349-5	Sequence 5, Appl
260	19	95.0	457	4	US-09-252-991A-28824	Sequence 28824, A	333	19	95.0	493	4	US-09-328-352-6753	Sequence 6753, Ap
261	19	95.0	467	4	US-09-543-681A-6871	Sequence 6871, Ap	334	19	95.0	493	4	US-09-427-770-70	Sequence 70, Appl
262	19	95.0	468	4	US-09-252-991A-18608	Sequence 18608, A	335	19	95.0	493	4	US-09-427-769-70	Sequence 70, Appl
263	19	95.0	469	4	US-09-543-681A-5423	Sequence 5423, Ap	336	19	95.0	493	4	US-09-540-236-2869	Sequence 2869, Ap
264	19	95.0	474	4	US-09-252-991A-27524	Sequence 27524, A	337	19	95.0	495	3	US-09-179-966-1	Sequence 1, Appl
265	19	95.0	475	4	US-09-252-991A-32806	Sequence 32806, A	338	19	95.0	495	3	US-08-980-994-1	Sequence 1, Appl
266	19	95.0	479	4	US-09-328-352-5396	Sequence 5396, Ap	339	19	95.0	504	4	US-09-252-991A-20317	Sequence 20317, A
267	19	95.0	485	2	US-08-446-803-1	Sequence 1, Appl	340	19	95.0	506	3	US-09-232-191-9	Sequence 9, Appl
268	19	95.0	485	2	US-08-446-803-2	Sequence 2, Appl	341	19	95.0	506	3	US-09-232-200-9	Sequence 9, Appl
269	19	95.0	485	2	US-08-861-837-1	Sequence 1, Appl	342	19	95.0	506	3	US-09-232-200-95	Sequence 95, Appl
270	19	95.0	485	2	US-08-861-837-2	Sequence 2, Appl	343	19	95.0	506	4	US-09-232-197-9	Sequence 9, Appl
271	19	95.0	485	2	US-08-600-908A-12	Sequence 12, Appl	344	19	95.0	506	4	US-09-232-197-95	Sequence 95, Appl
272	19	95.0	485	3	US-08-683-838A-12	Sequence 12, Appl	345	19	95.0	506	4	US-09-232-201-9	Sequence 9, Appl
273	19	95.0	485	3	US-08-600-656-1	Sequence 1, Appl	346	19	95.0	506	4	US-09-232-201-95	Sequence 95, Appl
274	19	95.0	485	3	US-08-600-656-2	Sequence 2, Appl	347	19	95.0	506	4	US-09-232-195-9	Sequence 95, Appl
275	19	95.0	485	3	US-08-600-656-7	Sequence 7, Appl	348	19	95.0	506	4	US-09-232-195-95	Sequence 95, Appl
276	19	95.0	485	3	US-09-170-670-1	Sequence 1, Appl	349	19	95.0	516	4	US-09-986-676A-2	Sequence 2, Appl
277	19	95.0	485	3	US-09-170-670-2	Sequence 2, Appl	350	19	95.0	522	4	US-09-252-991A-30451	Sequence 30451, A
278	19	95.0	485	3	US-09-170-670-6	Sequence 6, Appl	351	19	95.0	524	2	US-08-928-692-12	Sequence 12, Appl
279	19	95.0	485	3	US-09-170-670-7	Sequence 7, Appl	352	19	95.0	524	3	US-08-957-302A-2	Sequence 2, Appl
280	19	95.0	485	3	US-09-170-670-8	Sequence 8, Appl	353	19	95.0	524	3	US-09-342-403-2	Sequence 2, Appl
281	19	95.0	485	3	US-09-133-068-1	Sequence 1, Appl	354	19	95.0	532	4	US-09-359-372-12	Sequence 12, Appl
282	19	95.0	485	3	US-09-133-068-2	Sequence 2, Appl	355	19	95.0	532	4	US-09-252-991A-27288	Sequence 27288, A
283	19	95.0	485	3	US-09-133-068-6	Sequence 6, Appl	356	19	95.0	536	4	US-09-107-532A-6930	Sequence 6930, Ap
284	19	95.0	485	3	US-09-133-068-7	Sequence 7, Appl	357	19	95.0	538	4	US-09-489-039A-13518	Sequence 13518, A
285	19	95.0	485	3	US-09-133-068-8	Sequence 8, Appl	358	19	95.0	540	4	US-09-621-451-2	Sequence 2, Appl
286	19	95.0	485	3	US-09-183-412-1	Sequence 1, Appl	359	19	95.0	540	4	US-10-233-355-2	Sequence 2, Appl
287	19	95.0	485	3	US-09-183-412-2	Sequence 2, Appl	360	19	95.0	544	4	US-09-328-352-4446	Sequence 4446, Ap
288	19	95.0	485	3	US-09-183-412-6	Sequence 6, Appl	361	19	95.0	547	3	US-08-83-577A-149	Sequence 149, App
289	19	95.0	485	3	US-09-183-412-7	Sequence 7, Appl	362	19	95.0	547	3	US-08-897-438-149	Sequence 149, App
290	19	95.0	485	3	US-09-183-412-8	Sequence 8, Appl	363	19	95.0	547	4	US-08-649-518-149	Sequence 149, App
291	19	95.0	485	3	US-09-264-097-5	Sequence 5, Appl	364	19	95.0	550	4	US-09-396-478A-2	Sequence 2, Appl
292	19	95.0	485	3	US-09-264-097-7	Sequence 7, Appl	365	19	95.0	552	4	US-09-489-039A-13735	Sequence 13735, A
293	19	95.0	485	3	US-09-354-191A-1	Sequence 1, Appl	366	19	95.0	557	4	US-09-134-001C-5569	Sequence 5569, Ap
294	19	95.0	485	3	US-09-354-191A-2	Sequence 2, Appl	367	19	95.0	557	4	US-09-489-039A-13795	Sequence 13795, A
295	19	95.0	485	3	US-09-354-191A-7	Sequence 7, Appl	368	19	95.0	557	4	US-09-134-000C-4354	Sequence 4354, Ap
296	19	95.0	485	4	US-09-291-023A-13	Sequence 13, Appl	369	19	95.0	558	4	US-09-252-991A-27678	Sequence 27678, A
297	19	95.0	485	4	US-09-291-023A-18	Sequence 18, Appl	370	19	95.0	566	4	US-09-513-783A-142	Sequence 142, App
298	19	95.0	485	4	US-09-291-023A-19	Sequence 19, Appl	371	19	95.0	567	2	US-08-504-459-4	Sequence 4, Appl
299	19	95.0	485	4	US-09-290-734-1	Sequence 1, Appl	372	19	95.0	567	2	US-08-504-459-6	Sequence 6, Appl
300	19	95.0	485	4	US-09-290-734-2	Sequence 2, Appl	373	19	95.0	571	3	US-08-809-326A-6	Sequence 6, Appl
301	19	95.0	485	4	US-09-290-734-6	Sequence 6, Appl	374	19	95.0	571	4	US-09-689-914A-6	Sequence 6, Appl
302	19	95.0	485	4	US-09-290-734-7	Sequence 7, Appl	375	19	95.0	571	4	US-09-689-916A-6	Sequence 6, Appl
303	19	95.0	485	4	US-09-290-734-8	Sequence 8, Appl	376	19	95.0	571	4	US-09-689-916A-6	Sequence 6, Appl
304	19	95.0	485	4	US-09-290-734-24	Sequence 24, Appl	377	19	95.0	573	4	US-09-328-352-4675	Sequence 4675, Ap
305	19	95.0	485	4	US-09-290-734-26	Sequence 26, Appl	378	19	95.0	580	4	US-09-198-452A-332	Sequence 332, App
306	19	95.0	485	4	US-09-636-252A-12	Sequence 12, Appl	379	19	95.0	587	4	US-08-635-552A-3	Sequence 3, Appl
307	19	95.0	485	4	US-09-417-359A-5	Sequence 5, Appl	380	19	95.0	587	4	US-08-675-499A-4	Sequence 4, Appl
308	19	95.0	485	4	US-09-381-687-1	Sequence 1, Appl	381	19	95.0	587	4	US-08-812-008-4	Sequence 4, Appl
309	19	95.0	485	4	US-09-381-687-2	Sequence 2, Appl	382	19	95.0	589	4	US-09-643-657-14	Sequence 14, Appl
310	19	95.0	485	4	US-09-381-687-3	Sequence 3, Appl	383	19	95.0	591	2	US-08-736-770-5	Sequence 5, Appl
311	19	95.0	485	4	US-09-381-687-5	Sequence 5, Appl	384	19	95.0	591	2	US-09-643-657-5	Sequence 5, Appl
312	19	95.0	485	4	US-09-545-586-1	Sequence 1, Appl	385	19	95.0	591	4	US-09-643-657-15	Sequence 15, Appl
313	19	95.0	485	4	US-09-545-586-2	Sequence 2, Appl	386	19	95.0	597	4	US-09-252-991A-23152	Sequence 23152, A
314	19	95.0	485	4	US-09-545-586-6	Sequence 6, Appl	387	19	95.0	597	4	US-09-540-236-2805	Sequence 2805, Ap
315	19	95.0	485	4	US-09-545-586-7	Sequence 7, Appl	388	19	95.0	607	2	US-08-472-534-5	Sequence 5, Appl
316	19	95.0	485	4	US-09-545-586-8	Sequence 8, Appl	389	19	95.0	608	2	US-08-736-770-1	Sequence 1, Appl
317	19	95.0	485	4	US-09-545-586-24	Sequence 24, Appl	390	19	95.0	620	4	US-08-637-670-40	Sequence 40, Appl
318	19	95.0	485	4	US-09-545-586-26	Sequence 26, Appl	391	19	95.0	623	3	US-09-041-991A-6	Sequence 6, Appl
319	19	95.0	485	4	US-09-540-715A-13	Sequence 13, Appl	392	19	95.0	623	3		

393	19	95.0	623	4	US-09-608-533A-6	Sequence 6, Appli	466	19	95.0	651	3	US-08-993-170A-56	Sequence 56, Appl
394	19	95.0	625	4	US-09-661-322A-48	Sequence 48, Appl	467	19	95.0	651	3	US-08-993-170A-58	Sequence 58, Appl
395	19	95.0	626	4	US-09-252-991A-20268	Sequence 20268, A	468	19	95.0	651	3	US-08-993-775B-52	Sequence 52, Appl
396	19	95.0	631	1	US-08-487-890A-111	Sequence 111, App	469	19	95.0	651	3	US-08-993-775B-56	Sequence 56, Appl
397	19	95.0	631	2	US-08-478-435-111	Sequence 111, App	470	19	95.0	651	3	US-08-993-775B-58	Sequence 58, Appl
398	19	95.0	631	2	US-08-337-483-111	Sequence 111, App	471	19	95.0	651	4	US-08-427-770-52	Sequence 52, Appl
399	19	95.0	631	2	US-08-478-373-111	Sequence 111, App	472	19	95.0	651	4	US-08-427-770-56	Sequence 56, Appl
400	19	95.0	631	2	US-08-474-671-111	Sequence 111, App	473	19	95.0	651	4	US-08-427-770-58	Sequence 58, Appl
401	19	95.0	631	3	US-08-483-577A-111	Sequence 111, App	474	19	95.0	651	4	US-08-427-769-52	Sequence 52, Appl
402	19	95.0	631	3	US-08-897-438-111	Sequence 111, App	475	19	95.0	651	4	US-08-427-769-56	Sequence 56, Appl
403	19	95.0	631	4	US-08-637-654-111	Sequence 111, App	476	19	95.0	651	4	US-08-427-769-58	Sequence 58, Appl
404	19	95.0	631	4	US-08-649-518-111	Sequence 111, App	477	19	95.0	652	3	US-08-996-441B-2	Sequence 2, Appli
405	19	95.0	632	3	US-09-232-200-34	Sequence 34, Appl	478	19	95.0	652	3	US-08-996-441B-4	Sequence 4, Appli
406	19	95.0	632	3	US-09-232-200-35	Sequence 35, Appl	479	19	95.0	652	3	US-08-996-441B-6	Sequence 6, Appli
407	19	95.0	632	3	US-09-232-200-39	Sequence 39, Appl	480	19	95.0	652	3	US-08-996-441B-8	Sequence 8, Appli
408	19	95.0	632	4	US-09-232-197-34	Sequence 34, Appl	481	19	95.0	652	3	US-08-996-441B-10	Sequence 10, Appl
409	19	95.0	632	4	US-09-232-197-35	Sequence 35, Appl	482	19	95.0	652	3	US-08-996-441B-12	Sequence 12, Appl
410	19	95.0	632	4	US-09-232-197-39	Sequence 39, Appl	483	19	95.0	652	3	US-08-996-441B-14	Sequence 14, Appl
411	19	95.0	632	4	US-09-232-201-34	Sequence 34, Appl	484	19	95.0	652	3	US-08-996-441B-16	Sequence 16, Appl
412	19	95.0	632	4	US-09-232-201-35	Sequence 35, Appl	485	19	95.0	652	3	US-08-996-441B-18	Sequence 18, Appl
413	19	95.0	632	4	US-09-232-201-39	Sequence 39, Appl	486	19	95.0	652	3	US-08-996-441B-20	Sequence 20, Appl
414	19	95.0	632	4	US-09-232-195-34	Sequence 34, Appl	487	19	95.0	652	3	US-08-996-441B-22	Sequence 22, Appl
415	19	95.0	632	4	US-09-232-195-35	Sequence 35, Appl	488	19	95.0	652	3	US-08-996-441B-24	Sequence 24, Appl
416	19	95.0	632	4	US-09-232-195-39	Sequence 39, Appl	489	19	95.0	652	3	US-08-996-441B-26	Sequence 26, Appl
417	19	95.0	635	4	US-09-107-532A-7135	Sequence 7135, Ap	490	19	95.0	652	3	US-08-996-441B-28	Sequence 28, Appl
418	19	95.0	638	4	US-09-489-039A-8996	Sequence 8996, Ap	491	19	95.0	652	3	US-08-996-441B-30	Sequence 30, Appl
419	19	95.0	639	4	US-09-328-352-4379	Sequence 4379, Ap	492	19	95.0	652	3	US-08-996-441B-32	Sequence 32, Appl
420	19	95.0	641	4	US-09-543-681A-6258	Sequence 6258, A	493	19	95.0	652	3	US-08-996-441B-34	Sequence 34, Appl
421	19	95.0	641	4	US-09-232-200-27	Sequence 27, Appl	494	19	95.0	652	3	US-08-996-441B-36	Sequence 36, Appl
422	19	95.0	643	3	US-09-232-200-41	Sequence 41, Appl	495	19	95.0	652	3	US-08-996-441B-38	Sequence 38, Appl
423	19	95.0	643	3	US-09-232-200-42	Sequence 42, Appl	496	19	95.0	652	3	US-08-996-441B-40	Sequence 40, Appl
424	19	95.0	643	3	US-09-232-200-45	Sequence 45, Appl	497	19	95.0	652	3	US-08-996-441B-42	Sequence 42, Appl
425	19	95.0	643	3	US-09-232-200-53	Sequence 53, Appl	498	19	95.0	652	3	US-08-996-441B-44	Sequence 44, Appl
426	19	95.0	643	3	US-09-232-200-71	Sequence 71, Appl	499	19	95.0	652	3	US-08-996-441B-46	Sequence 46, Appl
427	19	95.0	643	4	US-09-232-197-27	Sequence 27, Appl	500	19	95.0	652	3	US-08-996-441B-48	Sequence 48, Appl
428	19	95.0	643	4	US-09-232-197-41	Sequence 41, Appl	501	19	95.0	652	3	US-08-996-441B-50	Sequence 50, Appl
429	19	95.0	643	4	US-09-232-197-42	Sequence 42, Appl	502	19	95.0	652	3	US-08-996-441B-54	Sequence 54, Appl
430	19	95.0	643	4	US-09-232-197-45	Sequence 45, Appl	503	19	95.0	652	3	US-08-996-441B-60	Sequence 60, Appl
431	19	95.0	643	4	US-09-232-197-53	Sequence 53, Appl	504	19	95.0	652	3	US-08-996-441B-62	Sequence 62, Appl
432	19	95.0	643	4	US-09-232-201-27	Sequence 27, Appl	505	19	95.0	652	3	US-08-996-441B-64	Sequence 64, Appl
433	19	95.0	643	4	US-09-232-201-41	Sequence 41, Appl	506	19	95.0	652	3	US-08-996-441B-66	Sequence 66, Appl
434	19	95.0	643	4	US-09-232-201-45	Sequence 45, Appl	507	19	95.0	652	3	US-08-996-441B-68	Sequence 68, Appl
435	19	95.0	643	4	US-09-232-201-53	Sequence 53, Appl	508	19	95.0	652	3	US-08-996-441B-98	Sequence 98, Appl
436	19	95.0	643	4	US-09-232-201-45	Sequence 45, Appl	509	19	95.0	652	3	US-08-996-441B-108	Sequence 108, App
437	19	95.0	643	4	US-09-232-201-53	Sequence 53, Appl	510	19	95.0	652	3	US-08-996-441B-110	Sequence 110, App
438	19	95.0	643	4	US-09-232-201-71	Sequence 71, Appl	511	19	95.0	652	3	US-08-996-441B-111	Sequence 111, App
439	19	95.0	643	4	US-09-343-681A-6345	Sequence 6345, Ap	512	19	95.0	652	3	US-08-993-722A-2	Sequence 2, Appli
440	19	95.0	643	4	US-09-232-195-27	Sequence 27, Appl	513	19	95.0	652	3	US-08-993-722A-6	Sequence 6, Appli
441	19	95.0	643	4	US-09-232-195-41	Sequence 41, Appl	514	19	95.0	652	3	US-08-993-722A-6	Sequence 8, Appli
442	19	95.0	643	4	US-09-232-195-42	Sequence 42, Appl	515	19	95.0	652	3	US-08-993-722A-10	Sequence 10, Appl
443	19	95.0	643	4	US-09-232-195-45	Sequence 45, Appl	516	19	95.0	652	3	US-08-993-722A-12	Sequence 12, Appl
444	19	95.0	643	4	US-09-232-195-53	Sequence 53, Appl	517	19	95.0	652	3	US-08-993-722A-14	Sequence 14, Appl
445	19	95.0	643	4	US-09-232-195-71	Sequence 71, Appl	518	19	95.0	652	3	US-08-993-722A-16	Sequence 16, Appl
446	19	95.0	643	4	US-08-487-890A-6	Sequence 6, Appli	519	19	95.0	652	3	US-08-993-722A-18	Sequence 18, Appl
447	19	95.0	644	1	US-08-478-435-6	Sequence 6, Appli	520	19	95.0	652	3	US-08-993-722A-20	Sequence 20, Appl
448	19	95.0	644	2	US-08-337-483-6	Sequence 6, Appli	521	19	95.0	652	3	US-08-993-722A-22	Sequence 22, Appl
449	19	95.0	644	2	US-08-478-373-6	Sequence 6, Appli	522	19	95.0	652	3	US-08-993-722A-24	Sequence 24, Appl
450	19	95.0	644	2	US-08-474-671-6	Sequence 6, Appli	523	19	95.0	652	3	US-08-993-722A-26	Sequence 26, Appl
451	19	95.0	644	3	US-08-483-577A-6	Sequence 6, Appli	524	19	95.0	652	3	US-08-993-722A-28	Sequence 28, Appl
452	19	95.0	644	3	US-08-483-577A-6	Sequence 6, Appli	525	19	95.0	652	3	US-08-993-722A-30	Sequence 30, Appl
453	19	95.0	644	4	US-08-483-577A-6	Sequence 6, Appli	526	19	95.0	652	3	US-08-993-722A-32	Sequence 32, Appl
454	19	95.0	644	4	US-08-483-577A-6	Sequence 6, Appli	527	19	95.0	652	3	US-08-993-722A-34	Sequence 34, Appl
455	19	95.0	644	4	US-08-483-577A-6	Sequence 6, Appli	528	19	95.0	652	3	US-08-993-722A-36	Sequence 36, Appl
456	19	95.0	647	3	US-08-483-577A-148	Sequence 148, App	529	19	95.0	652	3	US-08-993-722A-38	Sequence 38, Appl
457	19	95.0	647	3	US-08-897-438-148	Sequence 148, App	530	19	95.0	652	3	US-08-993-722A-40	Sequence 40, Appl
458	19	95.0	647	4	US-08-449-518-148	Sequence 148, App	531	19	95.0	652	3	US-08-993-722A-42	Sequence 42, Appl
459	19	95.0	651	3	US-08-996-441B-52	Sequence 52, Appl	532	19	95.0	652	3	US-08-993-722A-44	Sequence 44, Appl
460	19	95.0	651	3	US-08-996-441B-56	Sequence 56, Appl	533	19	95.0	652	3	US-08-993-722A-46	Sequence 46, Appl
461	19	95.0	651	3	US-08-996-441B-58	Sequence 58, Appl	534	19	95.0	652	3	US-08-993-722A-48	Sequence 48, Appl
462	19	95.0	651	3	US-08-993-722A-52	Sequence 52, Appl	535	19	95.0	652	3	US-08-993-722A-50	Sequence 50, Appl
463	19	95.0	651	3	US-08-993-722A-56	Sequence 56, Appl	536	19	95.0	652	3	US-08-993-722A-54	Sequence 54, Appl
464	19	95.0	651	3	US-08-993-722A-58	Sequence 58, Appl	537	19	95.0	652	3	US-08-993-722A-60	Sequence 60, Appl
465	19	95.0	651	3	US-08-993-170A-52	Sequence 52, Appl	538	19	95.0	652	3	US-08-993-722A-60	Sequence 60, Appl



685 19 95.0 652 4 US-09-427-769-68 Sequence 68, Appl 758  
686 19 95.0 652 4 US-09-427-769-98 Sequence 98, Appl 759  
687 19 95.0 652 4 US-09-427-769-100 Sequence 100, Appl 760  
688 19 95.0 652 4 US-09-427-769-108 Sequence 108, Appl 761  
689 19 95.0 652 4 US-09-427-769-110 Sequence 110, Appl 762  
690 19 95.0 652 4 US-09-427-769-111 Sequence 111, Appl 763  
691 19 95.0 652 5 PCT-US92-00040-2 Sequence 2, Appl 764  
692 19 95.0 652 6 5187091-2 Patent No. 5187091 765  
693 19 95.0 653 3 US-08-996-441B-100 Sequence 100, Appl 766  
694 19 95.0 653 3 US-08-993-722A-100 Sequence 100, Appl 767  
695 19 95.0 653 3 US-08-993-170A-100 Sequence 100, Appl 768  
696 19 95.0 653 3 US-08-993-775B-100 Sequence 100, Appl 769  
697 19 95.0 653 4 US-09-377-466B-8 Sequence 8, Appl 770  
698 19 95.0 653 4 US-09-377-466B-10 Sequence 10, Appl 771  
699 19 95.0 653 4 US-09-377-466B-12 Sequence 12, Appl 772  
700 19 95.0 653 4 US-09-377-466B-14 Sequence 14, Appl 773  
701 19 95.0 653 4 US-09-377-466B-16 Sequence 16, Appl 774  
702 19 95.0 653 4 US-09-377-466B-18 Sequence 18, Appl 775  
703 19 95.0 653 4 US-09-377-466B-20 Sequence 20, Appl 776  
704 19 95.0 653 4 US-09-377-466B-22 Sequence 22, Appl 777  
705 19 95.0 653 4 US-09-377-466B-24 Sequence 24, Appl 778  
706 19 95.0 653 4 US-09-377-466B-37 Sequence 37, Appl 779  
707 19 95.0 653 4 US-09-427-770-100 Sequence 100, Appl 780  
708 19 95.0 653 4 US-08-427-769-100 Sequence 100, Appl 781  
709 19 95.0 654 1 US-08-487-890A-12 Sequence 12, Appl 782  
710 19 95.0 654 2 US-08-478-435-12 Sequence 12, Appl 783  
711 19 95.0 654 2 US-08-337-483-12 Sequence 12, Appl 784  
712 19 95.0 654 3 US-08-478-373-12 Sequence 12, Appl 785  
713 19 95.0 654 3 US-08-474-671-12 Sequence 12, Appl 786  
714 19 95.0 654 3 US-08-483-577A-12 Sequence 12, Appl 787  
715 19 95.0 654 3 US-08-897-438-12 Sequence 12, Appl 788  
716 19 95.0 654 4 US-08-537-654-12 Sequence 12, Appl 789  
717 19 95.0 654 4 US-08-649-518-12 Sequence 12, Appl 790  
718 19 95.0 657 3 US-08-613-009A-19 Sequence 19, Appl 791  
719 19 95.0 657 4 US-08-778-570B-25 Sequence 25, Appl 792  
720 19 95.0 657 4 US-09-059-584-25 Sequence 25, Appl 793  
721 19 95.0 657 4 US-09-252-991A-27358 Sequence 27358, A 794  
722 19 95.0 660 1 US-08-487-890A-8 Sequence 8, Appl 795  
723 19 95.0 660 1 US-08-487-890A-10 Sequence 10, Appl 796  
724 19 95.0 660 2 US-08-478-435-8 Sequence 8, Appl 797  
725 19 95.0 660 2 US-08-478-435-10 Sequence 10, Appl 798  
726 19 95.0 660 2 US-08-337-483-8 Sequence 8, Appl 799  
727 19 95.0 660 2 US-08-337-483-10 Sequence 10, Appl 800  
728 19 95.0 660 2 US-08-478-373-8 Sequence 8, Appl 801  
729 19 95.0 660 2 US-08-478-373-10 Sequence 10, Appl 802  
730 19 95.0 660 3 US-08-474-671-10 Sequence 10, Appl 803  
731 19 95.0 660 3 US-08-483-577A-8 Sequence 8, Appl 804  
732 19 95.0 660 3 US-08-483-577A-10 Sequence 10, Appl 805  
733 19 95.0 660 3 US-08-837-438-8 Sequence 8, Appl 806  
734 19 95.0 660 3 US-08-837-438-10 Sequence 10, Appl 807  
735 19 95.0 660 3 US-08-837-438-12 Sequence 12, Appl 808  
736 19 95.0 660 4 US-08-637-654-8 Sequence 8, Appl 809  
737 19 95.0 660 4 US-08-637-654-10 Sequence 10, Appl 810  
738 19 95.0 660 4 US-08-649-518-8 Sequence 8, Appl 811  
739 19 95.0 660 4 US-08-649-518-10 Sequence 10, Appl 812  
740 19 95.0 678 1 US-08-282-141-2 Sequence 2, Appl 813  
741 19 95.0 678 1 US-08-435-434-2 Sequence 2, Appl 814  
742 19 95.0 678 1 US-08-435-436-2 Sequence 2, Appl 815  
743 19 95.0 678 2 US-08-438-863-2 Sequence 2, Appl 816  
744 19 95.0 678 2 US-08-438-864-2 Sequence 2, Appl 817  
745 19 95.0 678 3 US-08-438-862-2 Sequence 2, Appl 818  
746 19 95.0 678 3 US-08-628-747-2 Sequence 2, Appl 819  
747 19 95.0 678 3 US-08-402-253-2 Sequence 2, Appl 820  
748 19 95.0 678 3 US-08-443-866B-2 Sequence 2, Appl 821  
749 19 95.0 680 4 US-09-252-991A-26639 Sequence 26639, A 822  
750 19 95.0 682 4 US-09-252-991A-24919 Sequence 24919, A 823  
751 19 95.0 685 4 US-09-252-991A-26723 Sequence 26723, A 824  
752 19 95.0 696 6 5262177-4 Patent No. 5262177 825  
753 19 95.0 698 2 US-08-175-158A-2 Sequence 2, Appl 826  
754 19 95.0 703 4 US-09-252-991A-23905 Sequence 23905, A 827  
755 19 95.0 718 4 US-09-252-991A-33109 Sequence 33109, A 828  
756 19 95.0 720 4 US-09-252-991A-18618 Sequence 18618, A 829  
757 19 95.0 720 4 US-09-134-000C-5631 Sequence 5631, Ap 830

736 4 US-09-252-991A-27827 Sequence 27827, A  
748 4 US-09-252-991A-28449 Sequence 28449, A  
749 4 US-09-563-737-100 Sequence 100, Appl  
749 4 US-09-563-737-100 Sequence 100, Appl  
753 4 US-09-328-352-5412 Sequence 5412, Appl  
753 4 US-09-328-352-5412 Sequence 5412, Appl  
771 4 US-09-252-991A-17243 Sequence 17243, A  
778 2 US-08-907-166-6 Sequence 6, Appl  
788 4 US-09-391-340-6 Sequence 6, Appl  
788 4 US-09-391-340-6 Sequence 6, Appl  
789 4 US-09-390-234-16 Sequence 16, Appl  
789 4 US-09-603-311-16 Sequence 16, Appl  
791 4 US-09-107-532A-6065 Sequence 6065, Appl  
791 4 US-09-513-783A-4 Sequence 4, Appl  
812 4 US-09-513-783A-6 Sequence 6, Appl  
812 4 US-09-513-783A-6 Sequence 6, Appl  
816 4 US-09-489-039A-10520 Sequence 10520, A  
816 4 US-09-489-039A-10520 Sequence 10520, A  
819 4 US-09-651-656-15 Sequence 15, Appl  
819 4 US-09-651-656-15 Sequence 15, Appl  
821 4 US-09-620-865-24 Sequence 24, Appl  
821 3 US-09-428-869-15 Sequence 15, Appl  
843 3 US-09-029-267-20 Sequence 20, Appl  
843 3 US-09-029-267-20 Sequence 20, Appl  
846 4 US-09-328-352-6648 Sequence 6648, Appl  
846 4 US-09-328-352-6648 Sequence 6648, Appl  
848 4 US-09-489-039A-13773 Sequence 13773, A  
848 4 US-09-489-039A-13773 Sequence 13773, A  
849 4 US-09-252-991A-17953 Sequence 17953, A  
873 4 US-09-252-991A-30504 Sequence 30504, A  
873 4 US-09-252-991A-30504 Sequence 30504, A  
925 4 US-09-252-991A-27057 Sequence 27057, A  
925 4 US-09-252-991A-27057 Sequence 27057, A  
927 4 US-09-134-001C-4831 Sequence 4831, Appl  
944 4 US-09-328-352-4401 Sequence 4401, Appl  
944 4 US-09-328-352-4401 Sequence 4401, Appl  
972 3 US-08-335-844A-24 Sequence 24, Appl  
972 3 US-08-335-844A-24 Sequence 24, Appl  
972 3 US-09-129-366-24 Sequence 24, Appl  
975 4 US-09-543-681A-5755 Sequence 5755, Appl  
975 4 US-09-543-681A-5755 Sequence 5755, Appl  
982 4 US-09-023-909A-7 Sequence 7, Appl  
982 4 US-09-023-909A-7 Sequence 7, Appl  
990 4 US-09-252-991A-22792 Sequence 22792, A  
990 4 US-09-252-991A-22792 Sequence 22792, A  
993 4 US-09-894-998A-50 Sequence 50, Appl  
993 4 US-09-894-998A-50 Sequence 50, Appl  
1009 4 US-09-693-146-4 Sequence 4, Appl  
1009 4 US-09-693-146-4 Sequence 4, Appl  
1037 4 US-09-894-998A-54 Sequence 54, Appl  
1037 4 US-09-894-998A-54 Sequence 54, Appl  
1056 4 US-09-513-783A-32 Sequence 32, Appl  
1056 4 US-09-513-783A-32 Sequence 32, Appl  
1061 4 US-09-252-991A-23691 Sequence 23691, A  
1061 4 US-09-252-991A-23691 Sequence 23691, A  
1074 2 US-08-470-058-2 Sequence 2, Appl  
1074 2 US-08-470-058-2 Sequence 2, Appl  
1074 3 US-09-037-188-2 Sequence 2, Appl  
1074 3 US-09-037-188-2 Sequence 2, Appl  
1113 4 US-09-894-998A-51 Sequence 51, Appl  
1113 4 US-09-894-998A-51 Sequence 51, Appl  
1114 4 US-09-975-413A-12 Sequence 12, Appl  
1114 4 US-09-975-413A-12 Sequence 12, Appl  
1125 4 US-09-513-783A-152 Sequence 152, Appl  
1125 4 US-09-513-783A-152 Sequence 152, Appl  
1158 4 US-09-252-991A-19672 Sequence 19672, A  
1158 4 US-09-252-991A-19672 Sequence 19672, A  
1410 2 US-08-470-058-4 Sequence 4, Appl  
1410 2 US-08-470-058-4 Sequence 4, Appl  
1410 3 US-09-037-188-4 Sequence 4, Appl  
1410 3 US-09-037-188-4 Sequence 4, Appl  
1410 3 US-09-285-310-4 Sequence 4, Appl  
1410 3 US-09-285-310-4 Sequence 4, Appl  
1477 3 US-08-492-459-10 Sequence 10, Appl  
1477 3 US-08-492-459-10 Sequence 10, Appl  
1477 3 US-08-423-752-10 Sequence 10, Appl  
1477 3 US-08-423-752-10 Sequence 10, Appl  
1477 3 US-08-945-994-3 Sequence 3, Appl  
1477 3 US-08-945-994-3 Sequence 3, Appl  
1477 3 US-08-716-873-24 Sequence 24, Appl  
1477 3 US-08-716-873-24 Sequence 24, Appl  
1477 3 US-09-368-431-24 Sequence 24, Appl  
1477 3 US-09-368-431-24 Sequence 24, Appl  
1477 4 US-09-447-223-10 Sequence 10, Appl  
1477 4 US-09-447-223-10 Sequence 10, Appl  
1477 4 US-09-447-223-10 Sequence 10, Appl  
1477 4 US-09-447-223-10 Sequence 10, Appl  
1490 4 US-09-252-991A-28442 Sequence 28442, A  
1490 4 US-09-252-991A-28442 Sequence 28442, A  
1610 4 US-09-513-783A-22 Sequence 22, Appl  
1610 4 US-09-513-783A-22 Sequence 22, Appl  
2165 5 PCT-US95-12507-2 Sequence 2, Appl  
2165 5 PCT-US95-12507-2 Sequence 2, Appl  
2227 3 US-08-475-886-2 Sequence 2, Appl  
2227 3 US-08-475-886-2 Sequence 2, Appl  
2227 3 US-08-475-886-6 Sequence 6, Appl  
2227 3 US-08-475-886-6 Sequence 6, Appl  
2227 3 US-08-397-233-2 Sequence 2, Appl  
2227 3 US-08-397-233-2 Sequence 2, Appl  
2227 3 US-08-397-233-2 Sequence 2, Appl  
2227 3 US-08-397-233-2 Sequence 2, Appl  
2227 3 US-09-171-387-2 Sequence 2, Appl  
2227 3 US-09-171-387-2 Sequence 2, Appl  
2227 4 US-09-653-499-4 Sequence 4, Appl  
2227 4 US-09-653-499-4 Sequence 4, Appl  
2227 4 US-09-653-499-6 Sequence 6, Appl  
2227 4 US-09-653-499-6 Sequence 6, Appl  
2227 4 US-10-104-966-12 Sequence 12, Appl  
2227 4 US-10-104-966-12 Sequence 12, Appl  
2227 4 US-10-135-988-4 Sequence 4, Appl  
2227 4 US-10-135-988-4 Sequence 4, Appl  
2227 4 US-10-135-988-6 Sequence 6, Appl  
2227 4 US-10-135-988-6 Sequence 6, Appl  
2308 1 US-08-015-973-1 Sequence 1, Appl  
2308 1 US-08-015-973-1 Sequence 1, Appl  
2308 2 US-08-448-164-1 Sequence 1, Appl  
2308 2 US-08-448-164-1 Sequence 1, Appl  
2308 3 US-08-081-929-2 Sequence 2, Appl

831	19	95.0	2314	4	US-09-816-703A-2	Sequence 2, Appli	904	18	90.0	149	1	US-08-100-874-2	Sequence 2, Appli
832	19	95.0	2629	2	US-08-751-189-4	Sequence 4, Appli	905	18	90.0	149	3	US-08-963-409-4	Sequence 4, Appli
833	19	95.0	2629	2	US-09-060-836-4	Sequence 4, Appli	906	18	90.0	149	3	US-08-641-873-20	Sequence 20, Appl
834	19	95.0	2629	3	US-09-184-445-4	Sequence 4, Appli	907	18	90.0	152	3	US-08-963-409-5	Sequence 5, Appli
835	19	95.0	5588	3	US-09-036-987A-6	Sequence 6, Appli	908	18	90.0	152	4	US-09-252-991A-18700	Sequence 18700, A
836	19	95.0	5588	3	US-09-370-700-6	Sequence 6, Appli	909	18	90.0	152	4	US-05-252-991A-31418	Sequence 31418, A
837	19	95.0	5588	4	US-09-603-207-6	Sequence 6, Appli	910	18	90.0	157	4	US-05-134-001C-2991	Sequence 2991, Ap
838	19	95.0	8991	4	US-08-714-741-32	Sequence 32, Appl	911	18	90.0	158	1	US-08-611-107-4	Sequence 4, Appli
839	18	90.0	9	4	US-09-372-338-6	Sequence 6, Appli	912	18	90.0	158	2	US-08-422-560A-4	Sequence 4, Appli
840	18	90.0	9	4	US-10-126-752-6	Sequence 6, Appli	913	18	90.0	158	3	US-08-468-793-4	Sequence 4, Appli
841	18	90.0	10	2	US-08-482-228-120	Sequence 120, App	914	18	90.0	160	2	US-08-602-941-1	Sequence 1, Appli
842	18	90.0	10	3	US-08-482-528-120	Sequence 120, App	915	18	90.0	160	3	US-08-961-264-1	Sequence 1, Appli
843	18	90.0	11	2	US-08-934-222-99	Sequence 99, Appl	916	18	90.0	160	4	US-09-442-099A-1	Sequence 1, Appli
844	18	90.0	11	2	US-08-933-402-99	Sequence 99, Appl	917	18	90.0	160	4	US-09-612-342-1	Sequence 1, Appli
845	18	90.0	11	2	US-09-207-621-99	Sequence 99, Appl	918	18	90.0	160	4	US-09-612-421A-1	Sequence 1, Appli
846	18	90.0	11	2	US-08-533-818-99	Sequence 99, Appl	919	18	90.0	160	4	US-09-744-401-1	Sequence 1, Appli
847	18	90.0	11	3	US-09-231-787-99	Sequence 99, Appl	920	18	90.0	162	3	US-09-108-020-50	Sequence 50, Appl
848	18	90.0	11	3	US-08-934-224-99	Sequence 99, Appl	921	18	90.0	172	4	US-09-107-532A-5847	Sequence 5847, Ap
849	18	90.0	11	3	US-08-933-843-99	Sequence 99, Appl	922	18	90.0	179	4	US-09-634-238-296	Sequence 296, App
850	18	90.0	11	3	US-08-934-223-99	Sequence 99, Appl	923	18	90.0	184	4	US-09-489-039A-11491	Sequence 11491, A
851	18	90.0	11	3	US-09-413-432-99	Sequence 99, Appl	924	18	90.0	185	1	US-07-741-940-6	Sequence 6, Appli
852	18	90.0	19	4	US-08-833-281-10	Sequence 10, Appl	925	18	90.0	185	1	US-08-289-548A-6	Sequence 6, Appli
853	18	90.0	20	1	US-08-658-130-1	Sequence 1, Appli	926	18	90.0	185	1	US-08-452-654-6	Sequence 6, Appli
854	18	90.0	20	1	US-08-793-825-22	Sequence 22, Appl	927	18	90.0	185	1	US-08-452-655B-6	Sequence 6, Appli
855	18	90.0	20	4	US-09-092-000-5	Sequence 5, Appli	928	18	90.0	185	2	US-08-865-336-3	Sequence 3, Appli
856	18	90.0	20	4	US-09-280-068-22	Sequence 22, Appl	929	18	90.0	185	2	US-08-865-336-3	Sequence 3, Appli
857	18	90.0	20	4	US-08-833-281-12	Sequence 12, Appl	930	18	90.0	185	3	US-08-450-582-6	Sequence 6, Appli
858	18	90.0	20	4	US-09-481-593-18	Sequence 18, Appl	931	18	90.0	185	4	US-08-449-731-6	Sequence 6, Appli
859	18	90.0	20	5	PCT-US96-08723-1	Sequence 1, Appli	932	18	90.0	185	4	US-09-138-452A-858	Sequence 858 App
860	18	90.0	22	2	US-08-907-673-1	Sequence 1, Appli	933	18	90.0	187	4	US-09-134-001C-4819	Sequence 4819, Ap
861	18	90.0	22	3	US-08-803-397-6	Sequence 6, Appli	934	18	90.0	189	4	US-08-134-000C-4845	Sequence 2, Appli
862	18	90.0	22	4	US-09-425-597-6	Sequence 6, Appli	935	18	90.0	199	1	US-08-602-262-2	Sequence 2, Appli
863	18	90.0	22	5	PCT-US95-12502-6	Sequence 6, Appli	936	18	90.0	199	3	US-09-004-716-2	Sequence 2, Appli
864	18	90.0	27	4	US-09-033-780A-82	Sequence 82, Appl	937	18	90.0	199	4	US-09-477-510-2	Sequence 2, Appli
865	18	90.0	28	4	US-09-033-780A-81	Sequence 81, Appl	938	18	90.0	199	4	US-03-543-681A-5219	Sequence 5219, Ap
866	18	90.0	30	4	US-09-033-780A-104	Sequence 104, App	939	18	90.0	201	5	PCT-US94-02889-2	Sequence 5, Appli
867	18	90.0	38	4	US-09-033-780A-83	Sequence 83, Appl	940	18	90.0	201	1	US-07-929-380B-5	Sequence 5, Appli
868	18	90.0	38	4	US-09-033-780A-84	Sequence 84, Appl	941	18	90.0	205	4	US-09-107-532A-5214	Sequence 5214, Ap
869	18	90.0	39	2	US-08-889-291-15	Sequence 15, Appl	942	18	90.0	208	1	US-08-631-607-3	Sequence 3, Appli
870	18	90.0	39	3	US-09-098-244-15	Sequence 15, Appl	943	18	90.0	208	1	US-09-098-358B-3	Sequence 3, Appli
871	18	90.0	39	4	US-09-375-314-15	Sequence 15, Appl	944	18	90.0	210	1	US-07-741-940-4	Sequence 4, Appli
872	18	90.0	39	4	US-09-767-335-15	Sequence 15, Appl	945	18	90.0	210	1	US-08-289-548A-4	Sequence 4, Appli
873	18	90.0	43	4	US-09-039-780A-89	Sequence 89, Appl	946	18	90.0	210	1	US-08-452-654-4	Sequence 4, Appli
874	18	90.0	44	4	US-09-039-780A-88	Sequence 88, Appl	947	18	90.0	210	1	US-08-452-655B-4	Sequence 4, Appli
875	18	90.0	44	4	US-09-039-780A-89	Sequence 89, Appl	948	18	90.0	210	3	US-08-450-582-4	Sequence 4, Appli
876	18	90.0	54	4	US-09-039-780A-90	Sequence 90, Appl	949	18	90.0	210	4	US-08-449-731-4	Sequence 4, Appli
877	18	90.0	56	4	US-09-039-780A-93	Sequence 93, Appl	950	18	90.0	211	4	US-09-328-352-5256	Sequence 5256, Ap
878	18	90.0	60	4	US-09-134-000C-4258	Sequence 4258, Ap	951	18	90.0	214	4	US-09-795-926-16	Sequence 16, Appl
879	18	90.0	61	4	US-09-107-532A-4993	Sequence 4993, Ap	952	18	90.0	220	4	US-09-107-532A-6803	Sequence 6803, Ap
880	18	90.0	66	4	US-09-543-681A-5898	Sequence 5898, Ap	953	18	90.0	223	1	US-07-708-885B-3	Sequence 3, Appli
881	18	90.0	69	4	US-09-328-352-5719	Sequence 5719, Ap	954	18	90.0	223	1	US-07-714-386-3	Sequence 3, Appli
882	18	90.0	88	4	US-09-733-210-1423	Sequence 1423, Ap	955	18	90.0	223	1	US-07-708-888A-3	Sequence 3, Appli
883	18	90.0	90	4	US-09-393-634-76	Sequence 76, Appl	956	18	90.0	229	4	US-09-489-039A-9391	Sequence 9391, Ap
884	18	90.0	101	4	US-09-107-532A-6901	Sequence 6901, Ap	957	18	90.0	232	4	US-09-134-000C-3556	Sequence 3556, Ap
885	18	90.0	105	4	US-09-489-039A-10362	Sequence 10362, A	958	18	90.0	235	3	US-09-390-721-4	Sequence 4, Appli
886	18	90.0	113	4	US-09-614-912-186	Sequence 186, App	959	18	90.0	237	3	US-09-248-335-48	Sequence 48, Appl
887	18	90.0	116	4	US-09-489-039A-12512	Sequence 12512, A	960	18	90.0	243	4	US-09-134-000C-3733	Sequence 3733, Ap
888	18	90.0	118	4	US-09-540-236-3531	Sequence 3531, Ap	961	18	90.0	245	4	US-09-543-681A-5699	Sequence 5699, Ap
889	18	90.0	122	4	US-09-540-236-3531	Sequence 3531, Ap	962	18	90.0	247	4	US-09-328-352-7311	Sequence 7311, Ap
890	18	90.0	131	4	US-09-328-352-4806	Sequence 4806, Ap	963	18	90.0	249	4	US-09-252-991A-32233	Sequence 32233, A
891	18	90.0	132	4	US-09-134-000C-3642	Sequence 3642, Ap	964	18	90.0	251	4	US-09-648-004-20	Sequence 20, Appl
892	18	90.0	136	4	US-09-732-210-211	Sequence 211, App	965	18	90.0	251	4	US-09-286-981B-4	Sequence 4, Appli
893	18	90.0	142	1	US-07-951-715A-24	Sequence 24, Appl	966	18	90.0	252	4	US-09-489-039A-13590	Sequence 13590, A
894	18	90.0	142	1	US-08-459-448A-24	Sequence 24, Appl	967	18	90.0	254	4	US-09-138-452A-632	Sequence 632, App
895	18	90.0	142	3	US-08-459-595A-24	Sequence 24, Appl	968	18	90.0	257	2	US-08-637-759B-92	Sequence 92, Appl
896	18	90.0	142	3	US-08-459-504B-24	Sequence 24, Appl	969	18	90.0	257	3	US-08-871-355A-92	Sequence 92, Appl
897	18	90.0	142	3	US-08-459-444-24	Sequence 24, Appl	970	18	90.0	257	4	US-09-201-945-92	Sequence 9, Appl
898	18	90.0	142	3	US-09-547-422-24	Sequence 24, Appl	971	18	90.0	262	1	US-08-622-353-9	Sequence 9, Appli
899	18	90.0	143	3	US-08-946-329A-52	Sequence 52, Appl	972	18	90.0	262	2	US-08-622-352A-11	Sequence 11, Appl
900	18	90.0	143	3	US-09-328-352-7008	Sequence 7008, Ap	973	18	90.0	262	3	US-08-826-390-11	Sequence 11, Appl
901	18	90.0	143	4	US-09-134-000C-4544	Sequence 4544, Ap	974	18	90.0	265	4	US-09-795-926-5	Sequence 6, Appli
902	18	90.0	144	4	US-09-543-681A-5427	Sequence 5427, Ap	975	18	90.0	265	4	US-09-134-000C-5370	Sequence 5370, Ap
903	18	90.0	145	3	US-08-720-625-5	Sequence 5, Appli	976	18	90.0	267	4	US-09-543-681A-6389	Sequence 6389, Ap

977 18 90.0 268 4 US-09-489-039A-7310 Sequence 7310, Ap  
978 18 90.0 269 4 US-09-540-236-2149 Sequence 2149, Ap  
979 18 90.0 270 4 US-09-489-039A-9313 Sequence 9313, Ap  
980 18 90.0 273 4 US-09-489-039A-12815 Sequence 12815, A  
981 18 90.0 280 4 US-09-489-039A-23529 Sequence 23529, A  
982 18 90.0 281 3 US-09-053-702-2 Sequence 2, Appl  
983 18 90.0 281 4 US-09-540-236-2349 Sequence 2349, Ap  
984 18 90.0 284 4 US-09-372-422A-8 Sequence 8, Appl  
985 18 90.0 284 4 US-09-372-422A-8 Sequence 4, Appl  
986 18 90.0 285 4 US-09-314-701-30 Sequence 30, Appl  
987 18 90.0 286 4 US-09-489-039A-13264 Sequence 13264, A  
988 18 90.0 288 4 US-09-372-422A-12 Sequence 12, Appl  
989 18 90.0 288 4 US-09-372-422A-16 Sequence 16, Appl  
990 18 90.0 288 4 US-09-372-422A-18 Sequence 18, Appl  
991 18 90.0 288 4 US-09-372-448A-2 Sequence 2, Appl  
992 18 90.0 289 4 US-09-372-422A-2 Sequence 2, Appl  
993 18 90.0 289 4 US-09-372-422A-14 Sequence 14, Appl  
994 18 90.0 292 4 US-09-372-422A-4 Sequence 4, Appl  
995 18 90.0 292 4 US-09-372-422A-10 Sequence 10, Appl  
996 18 90.0 292 4 US-09-489-039A-9431 Sequence 9431, Ap  
997 18 90.0 295 4 US-09-134-000C-3576 Sequence 3576, Ap  
998 18 90.0 296 4 US-09-372-422A-20 Sequence 20, Appl  
999 18 90.0 296 4 US-09-328-352-6437 Sequence 6437, Ap  
1000 18 90.0 296 4 US-09-540-236-2843 Sequence 2843, Ap

## ALIGNMENTS

RESULT 1  
US-09-724-566A-78  
; Sequence 78, Application US/09724566A  
; Patent No. 6627739  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, John P.  
; APPLICANT: Basi, Gurigbal  
; APPLICANT: Doane, Minh Tam  
; APPLICANT: Frigon, No. 6627739mand  
; APPLICANT: John, Varghese  
; APPLICANT: Power, Michael  
; APPLICANT: Sinha, Sukanto  
; APPLICANT: Tatsuno, Gwen  
; APPLICANT: Tung, Jay  
; APPLICANT: Wang, Shuwen  
; APPLICANT: McConlogue, Lisa  
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and  
; FILE REFERENCE: 228-US-NEWC2  
; CURRENT APPLICATION NUMBER: US/09/724,566A  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 09/501,708  
; PRIOR FILING DATE: 2000-02-10  
; PRIOR APPLICATION NUMBER: 60/119,571  
; PRIOR FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 60/139,172  
; PRIOR FILING DATE: 1999-06-15  
; NUMBER OF SEQ ID NOS: 104  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 81  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide inhibitor P4-P4' XD-V  
; NAME/KEY: MOD\_RES  
; LOCATION: 3  
; OTHER INFORMATION: Xaa is hydroxyethylene or statine  
US-09-724-566A-81  
Query Match 95.0%; Score 19; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
Db 5 VAEF 8  
RESULT 3  
US-08-197-484-86  
; Sequence 86, Application US/08197484  
; Patent No. 641931  
; GENERAL INFORMATION:  
; APPLICANT: VITIELLO, Maria A.  
; APPLICANT: CHESTNUT, Robert W.  
; APPLICANT: SETTE, Alessandro D.  
; APPLICANT: CELIS, Eteban  
; APPLICANT: GRAY, Howard  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
; TITLE OF INVENTION: CTL IMMUNITY  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco

```
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/197,484
; FILING DATE: 16-FEB-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (206) 623-6793
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-197-484-86

Query Match 95.0%; Score 19; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
Db 2 VAEF 5

RESULT 4
; US-08-197-484-145
; Sequence 145, Application US/08197484
; Patent No. 6419931
; GENERAL INFORMATION:
; APPLICANT: VITIELLO, Maria A.
; APPLICANT: CHESTNUT, Robert W.
; APPLICANT: SETTE, Alessandro D.
; APPLICANT: CELIS, Esteban
; APPLICANT: GRAY, Howard
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/197,484
; FILING DATE: 16-FEB-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (206) 623-6793
; INFORMATION FOR SEQ ID NO: 145:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-197-484-145

Query Match 95.0%; Score 19; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
Db 2 VAEF 5

RESULT 5
; US-09-724-566A-73
; Sequence 73, Application US/09724566A
; Patent No. 6627739
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6627739mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; FILE OF INVENTION: Methods
; FILE REFERENCE: 228-US-NEWC2
; CURRENT APPLICATION NUMBER: US/09/724,566A
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/501,708
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 73
; LENGTH: 9
```



TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: P4-P4'staD-V peptide inhibitor  
NAME/KEY: MOD\_RES  
LOCATION: 5  
OTHER INFORMATION: Xaa is statine moiety  
US-09-724-566A-73

Query Match 95.0%; Score 19; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 6 VAEF 9

RESULT 6  
PCT-US93-02121-86  
; Sequence 86, Application PC/TUS9502121  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
; TITLE OF INVENTION: CTL IMMUNITY  
; NUMBER OF SEQUENCES: 153  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02121  
; FILING DATE: 16-FEB-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/197,484  
; FILING DATE: 16-FEB-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/935,811  
; FILING DATE: 26-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/874,491  
; FILING DATE: 27-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/827,682  
; FILING DATE: 29-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,568  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parmelee, Steven W.  
; REGISTRATION NUMBER: 31,990  
; REFERENCE/DOCKET NUMBER: 14137-26-4PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 467-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 86:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
PCT-US95-02121-86

Query Match 95.0%; Score 19; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 6 VAEF 9

Db 2 VAEF 5

RESULT 7  
PCT-US95-02121-145  
; Sequence 145, Application PC/TUS9502121  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
; TITLE OF INVENTION: CTL IMMUNITY  
; NUMBER OF SEQUENCES: 153  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02121  
; FILING DATE: 16-FEB-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/197,484  
; FILING DATE: 16-FEB-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/935,811  
; FILING DATE: 26-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/874,491  
; FILING DATE: 27-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/827,682  
; FILING DATE: 29-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,568  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parmelee, Steven W.  
; REGISTRATION NUMBER: 31,990  
; REFERENCE/DOCKET NUMBER: 14137-26-4PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 467-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 145:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
PCT-US95-02121-145

Query Match 95.0%; Score 19; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 2 VAEF 5

RESULT 8  
US-09-724-566A-72  
; Sequence 72, Application US/09724566A  
; Patent No. 6627739  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, John P.  
; APPLICANT: Basi, Gurigbal  
; APPLICANT: Doane, Minh Tam  
; APPLICANT: Frigon, No. 6627739mand  
; APPLICANT: John, Varghese  
; APPLICANT: Power, Michael  
; APPLICANT: Sinha, Sukanto  
; APPLICANT: Tatsuno, Gwen

APPLICANT: Tung, Jay  
APPLICANT: Wang, Shuwen  
APPLICANT: McConlogue, Lisa  
TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and  
TITLE OF INVENTION: Methods  
FILE REFERENCE: 228-US-NEWC2  
CURRENT APPLICATION NUMBER: US/09/724,566A  
CURRENT FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: US 09/501,708  
PRIOR FILING DATE: 2000-02-10  
PRIOR APPLICATION NUMBER: 60/119,571  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: 60/139,172  
PRIOR FILING DATE: 1999-06-15  
NUMBER OF SEQ ID NOS: 104  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 72  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: P10-P4'stad-V peptide inhibitor  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: 10  
OTHER INFORMATION: Xaa is statine moiety  
US-09-724-566A-72

Query Match 95.0%; Score 19; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 11 VAEF 14

RESULT 9  
US-09-724-566A-97  
Sequence 97, Application US/09724566A  
Patent No. 6627739  
GENERAL INFORMATION:  
APPLICANT: Anderson, John P.  
APPLICANT: Basi, Gurigbal  
APPLICANT: Doane, Minh Tam  
APPLICANT: Frigon, No. 6627739mand  
APPLICANT: John, Varghese  
APPLICANT: Power, Michael  
APPLICANT: Sinha, Sukanto  
APPLICANT: Tatsuno, Gwen  
APPLICANT: Tung, Jay  
APPLICANT: Wang, Shuwen  
APPLICANT: McConlogue, Lisa  
TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and  
TITLE OF INVENTION: Methods  
FILE REFERENCE: 228-US-NEWC2  
CURRENT APPLICATION NUMBER: US/09/724,566A  
CURRENT FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: US 09/501,708  
PRIOR FILING DATE: 2000-02-10  
PRIOR APPLICATION NUMBER: 60/119,571  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: 60/139,172  
PRIOR FILING DATE: 1999-06-15  
NUMBER OF SEQ ID NOS: 104  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 97  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: APP-derived fragment P10-P4' (D-V)  
US-09-724-566A-97

Query Match 95.0%; Score 19; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 11 VAEF 14

RESULT 10  
US-08-733-825-3  
Sequence 3, Application US/08733825  
Patent No. 5837839  
GENERAL INFORMATION:  
APPLICANT: Toth, Matthew J.  
APPLICANT: Huwylar, Leslie R.  
TITLE OF INVENTION: Coding Sequences for Mevalonate  
TITLE OF INVENTION: Pyrophosphate Decarboxylase  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 5837839artis Patent and Trademark Department  
STREET: 59 Route 10  
CITY: East Hanover  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07936-1080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/733,825  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/005,652  
FILING DATE: 18-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5837839ak, Henry P.  
REGISTRATION NUMBER: 33200  
REFERENCE/DOCKET NUMBER: 4-20615/P1/CGC 1834  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 277-5110  
TELEFAX: (908) 277-4606  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-733-825-3

Query Match 95.0%; Score 19; DB 2; Length 28;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 20 VAEF 23

RESULT 11  
US-08-487-890A-69  
Sequence 69, Application US/08487890A  
Patent No. 5708149  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele

APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Mardin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,890A  
FILING DATE: 07-JUN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 69:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-487-890A-69  
Query Match 95.0%; Score 19; DB 1; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;  
Qy 2 VAEF 5  
Db 14 VAEF 17  
RESULT 12  
US-08-468-763-1  
Sequence 1, Application US/08468763  
Patent No. 5741671  
GENERAL INFORMATION:  
APPLICANT: Agre, Peter C.  
TITLE OF INVENTION: Isolation, Cloning and Expression of  
TITLE OF INVENTION: Transmembrane Water Channel Proteins  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Allegretti  
STREET: 1001 G Street, N.W.  
CITY: Washington, D.C.  
STATE: D.C.  
COUNTRY: US  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,996A  
FILING DATE: 24-FEB-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Posorske, Laurence H.  
REGISTRATION NUMBER: 34,698  
REFERENCE/DOCKET NUMBER: 1107.48633  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 508-9100  
TELEFAX: 202 508-9299  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-393-996A-1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,763  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/393,996  
FILING DATE: 24-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Posorske, Laurence H.  
REGISTRATION NUMBER: 34,698  
REFERENCE/DOCKET NUMBER: 1107.48633  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 508-9100  
TELEFAX: 202 508-9299  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-468-763-1  
Query Match 95.0%; Score 19; DB 1; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;  
Qy 2 VAEF 5  
Db 14 VAEF 17  
RESULT 13  
US-08-393-996A-1  
Sequence 1, Application US/08393996A  
Patent No. 5858702  
GENERAL INFORMATION:  
APPLICANT: Agre, Peter C.  
TITLE OF INVENTION: Isolation, Cloning and Expression of  
TITLE OF INVENTION: Transmembrane Water Channel Proteins  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Allegretti  
STREET: 1001 G Street, N.W.  
CITY: Washington, D.C.  
STATE: D.C.  
COUNTRY: US  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,996A  
FILING DATE: 24-FEB-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Posorske, Laurence H.  
REGISTRATION NUMBER: 34,698  
REFERENCE/DOCKET NUMBER: 1107.48633  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 508-9100  
TELEFAX: 202 508-9299  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-393-996A-1

Query Match 95.0%; Score 19; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

Qy 2 VAEF 5  
Db 14 VAEF 17

## RESULT 14

US-08-478-435-69  
; Sequence 69, Application US/08478435  
; Patent No. 5923323  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena  
; APPLICANT: Harkness, Robin  
; APPLICANT: Schryvers, Anthony  
; APPLICANT: Chong, Pele  
; APPLICANT: Gray-Owen, Scott  
; APPLICANT: Yang, Yan-Ping  
; APPLICANT: Mordin, Andrew  
; APPLICANT: Klein, Michel  
; TITLE OF INVENTION: Transferrin Receptor Genes  
; NUMBER OF SEQUENCES: 147  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,435  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/337,483  
; FILING DATE: 08-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/175,116  
; FILING DATE: 29-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/149,968  
; FILING DATE: 08-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-462 MIS:V9  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-478-435-69

Query Match 95.0%; Score 19; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

Qy 2 VAEF 5  
Db 14 VAEF 17

Query Match 95.0%; Score 19; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

Qy 2 VAEF 5  
Db 14 VAEF 17

RESULT 15  
US-08-337-483-69  
; Sequence 69, Application US/08337483  
; Patent No. 5923562  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena  
; APPLICANT: Harkness, Robin  
; APPLICANT: Schryvers, Anthony  
; APPLICANT: Chong, Pele  
; APPLICANT: Gray-Owen, Scott  
; APPLICANT: Yang, Yan-Ping  
; APPLICANT: Mordin, Andrew  
; APPLICANT: Klein, Michel  
; TITLE OF INVENTION: Transferrin Receptor Genes  
; NUMBER OF SEQUENCES: 147  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/337,483  
; FILING DATE: 08-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-410 MIS:jb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-337-483-69  
Query Match 95.0%; Score 19; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

Qy 2 VAEF 5  
Db 14 VAEF 17

## RESULT 16

US-08-478-373-69  
; Sequence 69, Application US/08478373  
; Patent No. 5923841  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena  
; APPLICANT: Harkness, Robin  
; APPLICANT: Schryvers, Anthony  
; APPLICANT: Chong, Pele  
; APPLICANT: Gray-Owen, Scott  
; APPLICANT: Yang, Yan-Ping  
; APPLICANT: Mordin, Andrew  
; APPLICANT: Klein, Michel  
; TITLE OF INVENTION: Transferrin Receptor Genes  
; NUMBER OF SEQUENCES: 147  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney

```

; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,373
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-463 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-478-373-69

Query Match 95.0%; Score 19; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5
Db 14 VAEF 17

RESULT 17
; US-08-474-671-69
; Sequence 69, Application US/08474671
; Patent No. 6008326
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-ping
; APPLICANT: Mordin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,577A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,671
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-465 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-474-671-69

Query Match 95.0%; Score 19; DB 3; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5
Db 14 VAEF 17

RESULT 18
; US-08-483-577A-69
; Sequence 69, Application US/08483577A
; Patent No. 6015688
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-ping
; APPLICANT: Mordin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 160
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,577A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483

```

; FILING DATE: 08-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-511
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-483-577A-69
;
; Query Match 95.0%; Score 19; DB 3; Length 35;
; Best Local Similarity 100.0%; Pred. No. 1.3e+02;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 2 VAEF 5
; DB 14 VAEF 17
;
; RESULT 19
; US-08-897-438-69
; Sequence 69, Application US/08897438
; Patent No. 6262016
; GENERAL INFORMATION:
; APPLICANT: Loesmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Mordin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 160
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/897,438
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,577
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483
; FILING DATE: 08-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968

; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-720
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-897-438-69
;
; Query Match 95.0%; Score 19; DB 3; Length 35;
; Best Local Similarity 100.0%; Pred. No. 1.3e+02;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 2 VAEF 5
; DB 14 VAEF 17
;
; RESULT 20
; US-08-637-654-69
; Sequence 69, Application US/08637654
; Patent No. 6358727
; GENERAL INFORMATION:
; APPLICANT: Loesmore, Sheena M
; APPLICANT: Harkness, Robin E
; APPLICANT: Schryvers, Anthony B
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Mordin, Andrew D
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: HAEMOPHILUS TRANSFERRIN RECEPTOR GENES
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,654
; FILING DATE: 05-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA94/00616
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-637-654-69

Query Match 95.0%; Score 19; DB 4; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 14 VAEF 17

## RESULT 21

US-08-649-518-69  
; Sequence 69, Application US/08649518  
; Patent No. 6361779  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena  
; APPLICANT: Harkness, Robin  
; APPLICANT: Schryvers, Anthony  
; APPLICANT: Chong, Pele  
; APPLICANT: Gray-Owen, Scott  
; APPLICANT: Yang, Yan-Ping  
; APPLICANT: Mordin, Andrew  
; APPLICANT: Klein, Michel  
; TITLE OF INVENTION: Transferrin Receptor Genes  
; NUMBER OF SEQUENCES: 160  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/649,518  
; FILING DATE: 17-MAY-1996  
; CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/337,483  
; FILING DATE: 08-NOV-1994

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/483,577  
; FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/175,116  
; FILING DATE: 29-DEC-1993

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/148,968  
; FILING DATE: 08-NOV-1993

ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-608  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-649-518-69

Query Match 95.0%; Score 19; DB 4; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5

Db 14 VAEF 17

## RESULT 22

US-09-393-634-80  
; Sequence 80, Application US/09393634  
; Patent No. 6558910  
; GENERAL INFORMATION:  
; APPLICANT: Zuker, Charles S.  
; APPLICANT: Adler, Jon Elliot  
; APPLICANT: Ryba, Nick  
; APPLICANT: Mueller, Ken  
; APPLICANT: Hoon, Mark  
; APPLICANT: The Regents of the University of California  
; APPLICANT: as represented by the Secretary of the  
; APPLICANT: Department of Health and Human Services  
; TITLE OF INVENTION: SF, a No. 6558910el Family of Taste Receptors  
; FILE REFERENCE: 023075-080000US  
; CURRENT APPLICATION NUMBER: US/09/393,634  
; CURRENT FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 92  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 80  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human GR24  
; US-09-393-634-80

Query Match 95.0%; Score 19; DB 4; Length 68;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5

Db 14 VAEF 17

## RESULT 23

US-09-489-039A-13555  
; Sequence 13555, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 13555  
; LENGTH: 69  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
; US-09-489-039A-13555

Query Match 95.0%; Score 19; DB 4; Length 69;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5

Db 5 VAEF 8

## RESULT 24

US-09-328-352-7111  
; Sequence 7111, Application US/09328352



Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 7111  
LENGTH: 71  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
JS-09-328-352-7111

Query Match 95.0%; Score 19; DB 4; Length 71;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 2 VAEF 5  
26 VAEF 29

RESULT 25  
US-08-321-071A-10  
Sequence 10, Application US/08321071A  
Patent No. 5672866  
GENERAL INFORMATION:  
APPLICANT: CHITTENDEN, Thomas D.  
TITLE OF INVENTION: APOPTOSIS RELATED PROTEIN Bcl-Y, AND METHODS  
TITLE OF INVENTION: OF USE THEREOF  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/321,071A  
FILING DATE: 11-OCT-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/10103  
FILING DATE: 09-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/287,427  
FILING DATE: 09-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, HENRY N.  
REGISTRATION NUMBER: 32,073  
REFERENCE/DOCKET NUMBER: 104322.121CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-942-8400  
TELEFAX: 202-942-8484  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 73 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-321-071A-10

Query Match 95.0%; Score 19; DB 1; Length 73;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 47 VAEF 50

RESULT 26  
US-08-894-139-10  
Sequence 10, Application US/08894139  
Patent No. 6448376  
GENERAL INFORMATION:  
APPLICANT: LA THANGUE, NICHOLAS B.  
APPLICANT: BERNARDS, RENE  
APPLICANT: HIJWANS, ELEANORE M.  
TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,139  
FILING DATE: 13-AUG-1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 620-22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 74 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-139-10

Query Match 95.0%; Score 19; DB 4; Length 74;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 23 VAEF 26

RESULT 27  
US-09-732-210-1412  
Sequence 1412, Application US/09732210  
Patent No. 6573361  
GENERAL INFORMATION:  
APPLICANT: Bunkers, Greg J.  
APPLICANT: Liang, Jihong  
APPLICANT: Mittanck, Cindy A.  
APPLICANT: Seale, Jeffrey W.  
APPLICANT: Wu, Yonnie S.  
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
FILE REFERENCE: 38-21(15036)B  
CURRENT APPLICATION NUMBER: US/09/732,210  
CURRENT FILING DATE: 2000-12-07  
PRIOR APPLICATION NUMBER: US 60/169,513  
PRIOR FILING DATE: 1999-12-07

; PRIOR APPLICATION NUMBER: US 60/169,340  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 1753  
; SEQ ID NO 1412  
; LENGTH: 88  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae  
US-09-732-210-1412

Query Match 95.0%; Score 19; DB 4; Length 88;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
|  
|  
|  
|  
Db 11 VAEF 14

RESULT 28  
US-07-612-674-12  
; Sequence 12, Application US/07612674  
; Patent No. 5658792  
; GENERAL INFORMATION:  
; APPLICANT: NUEL, MARK J.  
; APPLICANT: MCCLUNG, J. KEITH  
; APPLICANT: STEWART, DAVID A.  
; APPLICANT: DANNER, DAVID B.  
; TITLE OF INVENTION: AN ANTIPROLIFERATIVE PROTEIN  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1100 NEW YORK AVENUE, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/612,674  
; FILING DATE: 19901114  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SCOTT, WATSON T.  
; REGISTRATION NUMBER: 26,581  
; REFERENCE/DOCKET NUMBER: 5683/82332  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3027  
; TELEFAX: 202-822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 93 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-07-612-674-12

Query Match 95.0%; Score 19; DB 1; Length 93;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
|  
|  
|  
|  
Db 18 VAEF 21

RESULT 29  
US-09-543-681A-5816

; Sequence 5816, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRAB1  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 5816  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-5816

Query Match 95.0%; Score 19; DB 4; Length 98;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
|  
|  
|  
|  
Db 21 VAEF 24

RESULT 30  
US-08-464-517-8  
; Sequence 8, Application US/08464517  
; Patent No. 5869640  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/464,517  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 100 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

JS-08-464-517-8

Query Match 95.0%; Score 19; DB 2; Length 100;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
DB 13 VAEF 16

## RESULT 31

JS-08-246-361A-8  
; Sequence 8, Application US/08246361A  
; Patent No. 5998582

## GENERAL INFORMATION:

APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USSES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/246,361A  
FILING DATE: 19-MAY-1994

## CLASSIFICATION: 435

PRIOR APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992

## ATTORNEY/AGENT INFORMATION:

NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941

## INFORMATION FOR SEQ ID NO: 8:

LENGTH: 100 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

Query Match 95.0%; Score 19; DB 2; Length 100;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
DB 13 VAEF 16

## RESULT 32

US-08-463-772-8  
; Sequence 8, Application US/08463772  
; Patent No. 6066501

## GENERAL INFORMATION:

APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USSES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463,772  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: US 07/963,308  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992

## ATTORNEY/AGENT INFORMATION:

NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941

## INFORMATION FOR SEQ ID NO: 8:

LENGTH: 100 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

Query Match 95.0%; Score 19; DB 3; Length 100;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
DB 13 VAEF 16

## RESULT 33

PCT-US93-05000-8  
; Sequence 8, Application PC/TUS9305000  
; GENERAL INFORMATION:

## APPLICANT: MITOTIX

## TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto

## NUMBER OF SEQUENCES: 42

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02173

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/05000  
FILING DATE: 19930525  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/888,178  
FILING DATE: 26-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: CSHL91-02A  
TELEPHONE: 617-861-6240  
TELEFAX: 616-861-9540  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCT-US93-05000-8

Query Match 95.0%; Score 19; DB 5; Length 100;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 13 VAEF 16

RESULT 34  
US-08-580-988A-21  
Sequence 21, Application US/08580988A  
Patent No. 5856161  
GENERAL INFORMATION:  
APPLICANT: Aggarwal et al.  
TITLE OF INVENTION: Tumor Necrosis Factor  
TITLE OF INVENTION: Receptor-I-Associated Protein Kinase And Methods  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dr. Benjamin A. Adler  
STREET: 8011 Candle Lane  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 Mb floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh  
SOFTWARE: Microsoft Word for Macintosh  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/580,988A  
FILING DATE: January 3, 1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benjamin Aaron Adler, Ph.D., J.D.  
REGISTRATION NUMBER: 35,423  
REFERENCE/DOCKET NUMBER: D5721CIP2  
TELEPHONE: 713-777-2321  
TELEFAX: 713-777-6908  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 101 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE:

DESCRIPTION: protein  
HYPOTHETICAL: no  
ANTI-SENSE: no  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-08-580-988A-21

Query Match 95.0%; Score 19; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 70 VAEF 73

RESULT 35  
US-09-540-236-3612  
Sequence 3612, Application US/09540236  
Patent No. 6673910  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAT?  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540,236  
CURRENT FILING DATE: 2000-04-04  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 3612  
LENGTH: 103  
TYPE: PRT  
ORGANISM: M.catarrhalis  
US-09-540-236-3612

Query Match 95.0%; Score 19; DB 4; Length 103;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 65 VAEF 68

RESULT 36  
US-08-464-517-25  
Sequence 25, Application US/08464517  
Patent No. 5869640  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,517  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
S-08-464-517-25

Query Match 95.0%; Score 19; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 4e+02; 0; Indels 0;  
Matches 4; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 VAEF 5  
Db 66 VAEF 69

RESULT 37  
US-08-246-361A-25  
Sequence 25, Application US/08246361A  
Patent No. 5998582  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THEREO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,361A  
FILING DATE: 19-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
US-08-246-361A-25  
Query Match 95.0%; Score 19; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 4e+02; 0; Indels 0;  
Matches 4; Conservative 0; Mismatches 0; Gaps 0;  
Qy 2 VAEF 5  
Db 66 VAEF 69  
RESULT 38  
US-08-463-772-25  
Sequence 25, Application US/08463772  
Patent No. 6066501  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THEREO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,772  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-463-772-25

Query Match 95.0%; Score 19; DB 3; Length 106;  
Best Local Similarity 100.0%; Pred. No. 4e+02; 0; Indels 0;  
Matches 4; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 VAEF 5  
Db 66 VAEF 69

RESULT 39  
US-09-489-039A-8721  
Sequence 8721, Application US/09489039A  
Patent No. 6610836

```

; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8721
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8721

Query Match          95.0%; Score 19; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 VAEF 5
Db      43 VAEF 46

RESULT 40
PCT-US93-05000-25
; Sequence 25, Application PC/TUS9305000
; GENERAL INFORMATION:
; APPLICANT: MITOXIX
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESS: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
PCT-US93-05000-25

Query Match          95.0%; Score 19; DB 5; Length 106;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 VAEF 5
Db      66 VAEF 69

```

```

RESULT 41
US-09-489-039A-9915
; Sequence 9915, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9915
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9915

Query Match          95.0%; Score 19; DB 4; Length 115;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 VAEF 5
Db      50 VAEF 53

RESULT 42
US-09-252-991A-21614
; Sequence 21614, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21614
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21614

Query Match          95.0%; Score 19; DB 4; Length 122;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 VAEF 5
Db      85 VAEF 88

RESULT 43
US-09-489-039A-7200
; Sequence 7200, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27

```

PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 7200  
LENGTH: 123  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-7200

Query Match 95.0%; Score 19; DB 4; Length 123;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
b 61 VAEF 64

RESULT 44  
US-08-529-055-53  
Sequence 53, Application US/08529055  
Patent No. 6592876  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: McDaniel, Larry S.  
APPLICANT: Swialo, Edwin  
APPLICANT: Yotter, Janet  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: Pneumococcal Genes, Portions  
TITLE OF INVENTION: Thereof, Expression Products  
TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
TITLE OF INVENTION: Portions and Products  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/529,055  
FILING DATE: 15-SEP-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2400  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 53:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 129 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-529-055-53

Query Match 95.0%; Score 19; DB 4; Length 129;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
b 7 VAEF 10

us-09-594-978a-3.ra

RESULT 45  
US-09-489-039A-12835  
Sequence 12835, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 12835  
LENGTH: 143  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-12835

Query Match 95.0%; Score 19; DB 4; Length 143;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 138 VAEF 141

RESULT 46  
US-09-198-452A-167  
Sequence 167, Application US/09198452A  
Patent No. 6559294  
GENERAL INFORMATION:  
APPLICANT: Grifais, R.  
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
TITLE OF INVENTION: and treatment of infection  
FILE REFERENCE: 9710-003-999  
CURRENT APPLICATION NUMBER: US/09/198,452A  
CURRENT FILING DATE: 1998-11-24  
NUMBER OF SEQ ID NOS: 6849  
SEQ ID NO 167  
LENGTH: 145  
TYPE: PRT  
ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-167

Query Match 95.0%; Score 19; DB 4; Length 145;  
Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 31 VAEF 34

RESULT 47  
US-09-134-000C-4292  
Sequence 4292, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: PatentIn version 3.1



SEQ ID NO 4292  
LENGTH: 145  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-4292

Query Match 95.0%; Score 19; DB 4; Length 145;  
Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 113 VAEF 116

## RESULT 48

US-08-460-694-3  
Sequence 3, Application US/08460694  
Patent No. 5858655  
GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: PRADI Cyclin and its cDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,694  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070002  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 150 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-460-694-3  
Query Match 95.0%; Score 19; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 5.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 66 VAEF 69

## RESULT 49

US-08-460-744-3  
Sequence 3, Application US/08460744  
Patent No. 6107541  
GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: PRADI Cyclin and its cDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,744  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070005  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 150 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-744-3

Query Match 95.0%; Score 19; DB 3; Length 150;  
Best Local Similarity 100.0%; Pred. No. 5.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 66 VAEF 69

## RESULT 50

US-07-667-711B-3  
Sequence 3, Application US/07667711B  
Patent No. 6110700  
GENERAL INFORMATION:  
APPLICANT: ARNOLD, ANDREW  
TITLE OF INVENTION: Pradi Cyclin and its cDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/667,711B  
FILING DATE: 11-MAR-1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MCPHAIL, DONALD R.  
REGISTRATION NUMBER: 35,811  
REFERENCE/DOCKET NUMBER: 0609.4070000  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:

US-08-460-694-3  
Query Match 95.0%; Score 19; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 5.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 66 VAEF 69

## RESULT 49

US-08-460-744-3  
Sequence 3, Application US/08460744  
Patent No. 6107541  
GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: PRADI Cyclin and its cDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:

```
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-07-667-711B-3

Query Match          95.0%; Score 19; DB 3; Length 150;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 VAEF 5
Db      66 VAEF 69

RESULT 51
US-08-679-493A-208
; Sequence 208, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 208
; LENGTH: 152
; TYPE: PRT
; ORGANISM: blueshark
US-08-679-493A-208

Query Match          95.0%; Score 19; DB 4; Length 152;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 VAEF 5
Db      92 VAEF 95

RESULT 52
US-09-134-000C-4087
; Sequence 4087, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4087
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4087

Query Match          95.0%; Score 19; DB 4; Length 153;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 VAEF 5

; LENGTH: 101 VAEF 104

RESULT 53
US-09-252-991A-20805
; Sequence 20805, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196,136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20805
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20805

Query Match          95.0%; Score 19; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 VAEF 5
Db      84 VAEF 87

RESULT 54
US-09-370-838-203
; Sequence 203, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 203
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-370-838-203

Query Match          95.0%; Score 19; DB 4; Length 164;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 VAEF 5
Db      100 VAEF 103

RESULT 55
US-09-540-236-2222
; Sequence 2222, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
```

APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540,236  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 2222  
LENGTH: 166  
TYPE: PRT  
ORGANISM: M.catarrhalis  
US-09-540-236-2222

Query Match 95.0%; Score 19; DB 4; Length 166;  
Best Local Similarity 100.0%; Pred. No. 6.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 97 VAEF 100

RESULT 56  
US-08-471-058-16  
Sequence 16, Application US/08471058  
Patent No. 5770443  
GENERAL INFORMATION:  
APPLICANT: Kiefer, Michael C.  
APPLICANT: Barr, Philip J.  
TITLE OF INVENTION: NOVEL APOPTOSIS MODULATING  
TITLE OF INVENTION: PROTEINS, DNA ENCODING THE PROTEINS AND METHODS OF USE  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,058  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/330,157  
FILING DATE: 07-OCT-1994  
APPLICATION NUMBER: 08/160,067  
FILING DATE: 30-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lehnardt, Susan K  
REGISTRATION NUMBER: 33,943  
REFERENCE/DOCKET NUMBER: 23647-200007.12  
TELEPHONE: 415-813-5600  
TELEFAX: 415-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 172 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-471-058-16

Query Match 95.0%; Score 19; DB 1; Length 172;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

us-09-594-978a-3.ra1

APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540,236  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 2222  
LENGTH: 166  
TYPE: PRT  
ORGANISM: M.catarrhalis  
US-09-540-236-2222

Query Match 95.0%; Score 19; DB 4; Length 166;  
Best Local Similarity 100.0%; Pred. No. 6.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 97 VAEF 100

RESULT 56  
US-08-471-058-16  
Sequence 16, Application US/08471058  
Patent No. 5770443  
GENERAL INFORMATION:  
APPLICANT: Kiefer, Michael C.  
APPLICANT: Barr, Philip J.  
TITLE OF INVENTION: NOVEL APOPTOSIS MODULATING  
TITLE OF INVENTION: PROTEINS, DNA ENCODING THE PROTEINS AND METHODS OF USE  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,058  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/330,157  
FILING DATE: 07-OCT-1994  
APPLICATION NUMBER: 08/160,067  
FILING DATE: 30-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lehnardt, Susan K  
REGISTRATION NUMBER: 33,943  
REFERENCE/DOCKET NUMBER: 23647-200007.12  
TELEPHONE: 415-813-5600  
TELEFAX: 415-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 172 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-471-058-16

Query Match 95.0%; Score 19; DB 1; Length 172;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 122 VAEF 125

RESULT 57  
US-08-690-095-3  
Sequence 3, Application US/08690095  
Patent No. 5792648  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Au-Young, Janice  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN MACROPHAGE ANTIGEN  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/690,095  
FILING DATE: Filed Herewith  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0110 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 172 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 293274  
US-08-690-095-3

Query Match 95.0%; Score 19; DB 1; Length 172;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 122 VAEF 125

RESULT 58  
US-08-471-057-16  
Sequence 16, Application US/08471057  
Patent No. 6015687  
GENERAL INFORMATION:  
APPLICANT: KIEFER, MICHAEL C.  
APPLICANT: BARR, PHILIP J.  
TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA  
TITLE OF INVENTION: ENCODING THE PROTEINS AND METHODS OF USE THEREOF  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto

STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,057  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/320,157  
FILING DATE: 07-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: LEHNHARDT, SUSAN K.  
REGISTRATION NUMBER: 33,943  
REFERENCE/DOCKET NUMBER: 23647-20007.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 172 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-471-057-16

Query Match 95.0%; Score 19; DB 3; Length 172;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 122 VAEF 125

RESULT 59  
US-09-113-789-3  
Sequence 3, Application US/09113789  
Patent No. 6034219  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Au-Yang, Janice  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN MACROPHAGE ANTIGEN  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/113,789  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/690,095  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0110 US  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 172 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 293274  
US-09-113-789-3

Query Match 95.0%; Score 19; DB 3; Length 172;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 122 VAEF 125

RESULT 60  
US-08-470-865-16  
Sequence 16, Application US/08470865  
Patent No. 6586395  
GENERAL INFORMATION:  
APPLICANT: KIEFER, MICHAEL C.  
APPLICANT: BARR, PHILIP J.  
TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA  
TITLE OF INVENTION: ENCODING THE PROTEINS AND METHODS OF USE THEREOF  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,865  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/320,157  
FILING DATE: 07-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: LEHNHARDT, SUSAN K.  
REGISTRATION NUMBER: 33,943  
REFERENCE/DOCKET NUMBER: 23647-20007.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 172 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-470-865-16

Query Match 95.0%; Score 19; DB 4; Length 172;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

ORGANISM: Klebsiella pneumoniae	US-09-489-039A-9869	Query Match	95.0%; Score 19; DB 4; Length 174;	Best Local Similarity	100.0%; Pred. No. 6.8e+02;	Mismatches	0; Conservative	0; Indels	0; Gaps	0;
Db	122 VAEF 125									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									
Mismatches	0; Conservative									
Indels	0; Gaps									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									
Mismatches	0; Conservative									
Indels	0; Gaps									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									
Mismatches	0; Conservative									
Indels	0; Gaps									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									
Mismatches	0; Conservative									
Indels	0; Gaps									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									
Mismatches	0; Conservative									
Indels	0; Gaps									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									
Mismatches	0; Conservative									
Indels	0; Gaps									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									
Mismatches	0; Conservative									
Indels	0; Gaps									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									
Mismatches	0; Conservative									
Indels	0; Gaps									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									
Mismatches	0; Conservative									
Indels	0; Gaps									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									
Mismatches	0; Conservative									
Indels	0; Gaps									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									
Mismatches	0; Conservative									
Indels	0; Gaps									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									
Mismatches	0; Conservative									
Indels	0; Gaps									
Query Match	95.0%; Score 19; DB 4; Length 174;									
Best Local Similarity	100.0%; Pred. No. 6.8e+02;									

## RESULT 65

JS-09-252-991A-24341  
; Sequence 24341, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24341  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
JS-09-252-991A-24341

Query Match 95.0%; Score 19; DB 4; Length 185;  
Best Local Similarity 100.0%; Pred. No. 7.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 2 VAEF 5

Db 4 VAEF 7

## RESULT 66

JS-09-543-681A-7313  
; Sequence 7313, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543.681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 7313  
; LENGTH: 189  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-7313

Query Match 95.0%; Score 19; DB 4; Length 189;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

Db 73 VAEF 76

## RESULT 67

US-08-816-241-1  
; Sequence 1, Application US/08816241  
; Patent No. 5804185  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Goli, Surya K.  
; TITLE OF INVENTION: NOVEL RNA EDITING ENZYME  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ For Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/816,241  
FILING DATE: Filed Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0239 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 190 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PROSTUT09  
CLONE: 1646823  
US-08-816-241-1

Query Match 95.0%; Score 19; DB 1; Length 190;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

Db 104 VAEF 107

## RESULT 68

US-09-128-395-1  
; Sequence 1, Application US/09128395  
; Patent No. 6087108  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Goli, Surya K.  
; TITLE OF INVENTION: NOVEL RNA EDITING ENZYME  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ For Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/128,395  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/816,241  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0239 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 190 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: PROSTUT09  
 ; CLONE: 1646823  
 ; US-09-128-395-1

Query Match 95.0%; Score 19; DB 3; Length 190;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 Db 104 VAEF 107

## RESULT 69

; US-09-134-000C-3800  
 ; Sequence 3800, Application US/09134000C  
 ; Patent No. 6617156  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 032796-032  
 ; CURRENT APPLICATION NUMBER: US/09/134,000C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/055,778  
 ; PRIOR FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 6812  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 3800  
 ; LENGTH: 194  
 ; TYPE: PRT  
 ; ORGANISM: Enterococcus faecalis  
 ; US-09-134-000C-3800

Query Match 95.0%; Score 19; DB 4; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 Db 115 VAEF 118

## RESULT 70

; US-09-232-200-85  
 ; Sequence 85, Application US/09232200A  
 ; Patent No. 6288213  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stahl, Andreas  
 ; APPLICANT: Hirsch, David J.  
 ; APPLICANT: Lodish, Harvey F.  
 ; APPLICANT: Gimeno, Ruth E.  
 ; APPLICANT: Tartaglia, Louis A.  
 ; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
 ; FILE REFERENCE: WH197-21P3MB  
 ; CURRENT APPLICATION NUMBER: US/09/232,200A  
 ; CURRENT FILING DATE: 1999-01-14  
 ; EARLIER APPLICATION NUMBER: 60/071,374  
 ; EARLIER FILING DATE: 1998-01-15  
 ; EARLIER APPLICATION NUMBER: 60/093,491

; EARLIER FILING DATE: 1998-07-20  
 ; EARLIER APPLICATION NUMBER: 60/110,941  
 ; EARLIER FILING DATE: 1998-12-04  
 ; NUMBER OF SEQ ID NOS: 105  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 85  
 ; LENGTH: 199  
 ; TYPE: PRT  
 ; ORGANISM: Aspergillus nidulans  
 ; US-09-232-200-85

Query Match 95.0%; Score 19; DB 3; Length 199;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 Db 99 VAEF 102

## RESULT 71

; US-09-232-197-85  
 ; Sequence 85, Application US/09232197A  
 ; Patent No. 6300096  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stahl, Andreas  
 ; APPLICANT: Hirsch, David J.  
 ; APPLICANT: Lodish, Harvey F.  
 ; APPLICANT: Gimeno, Ruth E.  
 ; APPLICANT: Tartaglia, Louis A.  
 ; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
 ; FILE REFERENCE: WH197-21P3MA  
 ; CURRENT APPLICATION NUMBER: US/09/232,197A  
 ; CURRENT FILING DATE: 1999-01-14  
 ; EARLIER APPLICATION NUMBER: 60/071,374  
 ; EARLIER FILING DATE: 1998-01-15  
 ; EARLIER APPLICATION NUMBER: 60/093,491  
 ; EARLIER FILING DATE: 1998-07-20  
 ; EARLIER APPLICATION NUMBER: 60/110,941  
 ; EARLIER FILING DATE: 1998-12-04  
 ; NUMBER OF SEQ ID NOS: 105  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 85  
 ; LENGTH: 199  
 ; TYPE: PRT  
 ; ORGANISM: Aspergillus nidulans  
 ; US-09-232-197-85

Query Match 95.0%; Score 19; DB 4; Length 199;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 Db 99 VAEF 102

## RESULT 72

; US-09-232-201-85  
 ; Sequence 85, Application US/09232201A  
 ; Patent No. 6348321  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stahl, Andreas  
 ; APPLICANT: Hirsch, David J.  
 ; APPLICANT: Lodish, Harvey F.  
 ; APPLICANT: Gimeno, Ruth E.  
 ; APPLICANT: Tartaglia, Louis A.  
 ; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
 ; FILE REFERENCE: WH197-21P3MC  
 ; CURRENT APPLICATION NUMBER: US/09/232,201A  
 ; CURRENT FILING DATE: 1999-01-14  
 ; EARLIER APPLICATION NUMBER: 60/071,374  
 ; EARLIER FILING DATE: 1998-01-15



EARLIER APPLICATION NUMBER: 60/093,491  
EARLIER FILING DATE: 1998-07-20  
EARLIER APPLICATION NUMBER: 60/110,941  
EARLIER FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 85  
LENGTH: 199  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
US-09-232-201-85

Query Match 95.0%; Score 19; DB 4; Length 199;  
Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 99 VAEF 102

RESULT 73  
US-09-232-195-85  
Sequence 85, Application US/09232195A  
Patent No. 6657049  
GENERAL INFORMATION:  
APPLICANT: Stahl, Andreas  
APPLICANT: Hirsch, David J.  
APPLICANT: Lodish, Harvey F.  
APPLICANT: Gimeno, Ruth E.  
APPLICANT: Tartaglia, Louis A.  
TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
FILE REFERENCE: WH197-21D3MD  
CURRENT APPLICATION NUMBER: US/09/232,195A  
CURRENT FILING DATE: 1998-01-04  
EARLIER APPLICATION NUMBER: 60/071,374  
EARLIER FILING DATE: 1998-01-15  
EARLIER APPLICATION NUMBER: 60/093,491  
EARLIER FILING DATE: 1998-07-20  
EARLIER APPLICATION NUMBER: 60/110,941  
EARLIER FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 85  
LENGTH: 199  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
US-09-232-195-85

Query Match 95.0%; Score 19; DB 4; Length 199;  
Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 99 VAEF 102

RESULT 74  
US-09-107-532A-4612  
Sequence 4612, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA

ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 4612:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 203 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...203  
SEQUENCE DESCRIPTION: SEQ ID NO: 4612:  
US-09-107-532A-4612

Query Match 95.0%; Score 19; DB 4; Length 203;  
Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 63 VAEF 66

RESULT 75  
US-08-684-024-1  
Sequence 1, Application US/08684024  
Patent No. 5834298  
GENERAL INFORMATION:  
APPLICANT: Benezra, Robert  
TITLE OF INVENTION: GENE ENCODING THE HUMAN HOMOLOG OF MAD2  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/684,024  
FILING DATE: 19-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 1747/46621-A

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 205 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-684-024-1

```

```

Query Match          95.0%; Score 19; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      2 VAEF 5
Db      20 VAEF 23

```

```

Search completed: May 24, 2004, 17:43:23
Job time : 16 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

M protein - protein search, using sw model

run on: May 24, 2004, 17:34:23 ; Search time 39.6429 Seconds  
(without alignments)  
35.179 Million cell updates/sec

title: US-09-594-978A-3  
effect score: 20  
sequence: 1 XVAEF 5

coring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

searched: 1149313 seqs, 278921704 residues

total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/prodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep.\*

4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*

7: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*

10: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*

11: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB.pep.\*

12: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep.\*

13: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6/prodata/1/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB.pep.\*

17: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

18: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	19	95.0	8	9	US-09-791-378-343	Sequence 343, App
2	19	95.0	8	10	US-09-791-393-206	Sequence 206, App
3	19	95.0	8	10	US-09-791-393-206	Sequence 206, App
4	19	95.0	9	14	US-10-128-711-86	Sequence 86, App
5	19	95.0	9	14	US-10-128-711-145	Sequence 145, App
6	19	95.0	10	12	US-09-908-943A-197	Sequence 197, App
7	19	95.0	12	12	US-09-908-943A-196	Sequence 196, App
8	19	95.0	12	14	US-10-032-818-34	Sequence 34, App
9	19	95.0	13	14	US-10-239-313A-379	Sequence 379, App
10	19	95.0	15	14	US-10-119-528-7	Sequence 7, Appli
11	19	95.0	20	10	US-09-171-432A-62	Sequence 62, Appl
12	19	95.0	20	10	US-09-171-432A-63	Sequence 63, Appl
13	19	95.0	25	10	US-09-171-432A-65	Sequence 65, Appl
14	19	95.0	28	9	US-09-864-761-47201	Sequence 47201, A
15	19	95.0	28	12	US-10-424-599-216297	Sequence 216297,

35	14	US-10-043-344-69	Sequence 69, Appl
37	15	US-10-242-355-570	Sequence 570, App
37	15	US-10-430-752A-49	Sequence 49, Appl
38	15	US-10-430-752A-45	Sequence 45, Appl
41	15	US-10-372-003A-71	Sequence 71, Appl
43	9	US-09-864-761-37202	Sequence 37202, A
43	12	US-10-424-599-183032	Sequence 183032,
45	12	US-10-424-599-171318	Sequence 171318,
48	9	US-09-864-761-38166	Sequence 38166, A
52	12	US-10-424-599-205032	Sequence 205032,
55	12	US-10-424-599-192247	Sequence 192247,
55	12	US-10-424-599-233502	Sequence 233502,
55	15	US-10-372-003A-63	Sequence 63, Appl
55	15	US-10-372-003A-67	Sequence 67, Appl
56	14	US-10-219-329-4	Sequence 4, Appli
56	14	US-10-153-185-4	Sequence 4, Appli
56	14	US-10-219-561-4	Sequence 4, Appli
57	12	US-10-424-599-215734	Sequence 215734,
57	12	US-10-424-599-260762	Sequence 260762,
57	15	US-10-372-003A-75	Sequence 75, Appl
58	12	US-10-424-599-214930	Sequence 214930,
60	9	US-09-864-761-45291	Sequence 45291, A
68	9	US-09-393-634-80	Sequence 80, Appl
68	11	US-09-864-408A-40	Sequence 40, Appl
68	12	US-10-364-861-80	Sequence 80, Appl
68	14	US-10-383-982-80	Sequence 80, Appl
69	9	US-09-864-761-45716	Sequence 45716, A
69	9	US-09-738-636-5596	Sequence 5596, Ap
69	12	US-10-424-599-185385	Sequence 185385,
70	14	US-10-029-386-28645	Sequence 28645, A
71	12	US-10-424-599-202605	Sequence 202605,
71	12	US-10-424-599-202605	Sequence 202605,
71	12	US-10-424-599-202605	Sequence 202605,
71	14	US-10-029-386-303481	Sequence 30348, A
72	14	US-10-029-386-303481	Sequence 30348, A
72	14	US-10-029-386-293320	Sequence 29320, A
74	14	US-10-214-188-10	Sequence 10, Appl
74	14	US-10-156-561-9071	Sequence 9071, Ap
75	12	US-10-424-599-156128	Sequence 156128,
75	12	US-10-424-599-261424	Sequence 261424,
76	12	US-10-424-599-253204	Sequence 253204,
77	12	US-10-424-599-211908	Sequence 211908,
77	12	US-10-335-977-7493	Sequence 7493, A
77	14	US-10-029-386-32709	Sequence 32709, A
78	12	US-10-424-599-162039	Sequence 162039,
78	12	US-10-424-599-262701	Sequence 262701,
79	12	US-10-363-616-368	Sequence 368, App
80	12	US-10-424-599-216097	Sequence 216097,
82	15	US-10-367-980A-14	Sequence 14, Appl
84	11	US-09-864-408A-5282	Sequence 5282, Ap
84	14	US-10-078-090-125	Sequence 125, App
85	12	US-10-424-599-155636	Sequence 155636,
85	12	US-10-424-599-201853	Sequence 201853,
89	9	US-09-815-242-11224	Sequence 11224, A
89	9	US-09-815-242-11245	Sequence 11245, A
89	12	US-10-282-122A-58489	Sequence 58489, A
89	12	US-10-282-122A-58530	Sequence 58530, A
89	12	US-10-282-122A-66960	Sequence 66960, A
89	12	US-10-282-122A-68974	Sequence 68974, A
89	12	US-10-424-599-180307	Sequence 180307,
89	14	US-10-097-111-360	Sequence 360, App
90	12	US-10-424-599-245830	Sequence 245830,
93	12	US-10-424-599-200533	Sequence 200533,
93	12	US-10-424-599-204369	Sequence 204369,
94	15	US-10-372-003A-79	Sequence 79, Appl
99	15	US-10-265-049-2775	Sequence 2775, Ap
100	12	US-10-335-977-7495	Sequence 7495, Ap
101	12	US-10-335-977-7494	Sequence 7494, Ap
103	12	US-10-424-599-196121	Sequence 196121,
104	12	US-10-424-599-245993	Sequence 245993,
104	16	US-10-294-445-30	Sequence 30, Appl
105	14	US-10-291-851-90	Sequence 90, Appl
107	9	US-09-738-626-4520	Sequence 4520, Ap
109	15	US-10-363-616-367	Sequence 367, App
109	15	US-10-264-237-1538	Sequence 1538, Ap

89	19	95.0	110	12	US-10-424-599-144932,	Sequence 144932,	162	19	95.0	189	12	US-10-424-599-209651	Sequence 209651,
90	19	95.0	110	12	US-10-424-599-257728,	Sequence 257728,	163	19	95.0	189	14	US-10-238-075-1037	Sequence 1037, Ap
91	19	95.0	111	12	US-10-424-599-192406,	Sequence 192406,	164	19	95.0	191	12	US-10-282-122A-77681	Sequence 77681, A
92	19	95.0	112	12	US-10-424-599-189066,	Sequence 189066,	165	19	95.0	192	12	US-10-424-599-269537	Sequence 269537,
93	19	95.0	115	12	US-10-424-599-222263,	Sequence 222263,	166	19	95.0	193	9	US-09-888-911-4	Sequence 4, Appl1
94	19	95.0	117	12	US-10-424-599-206883,	Sequence 206883,	167	19	95.0	193	12	US-10-425-114-43750	Sequence 43750, A
95	19	95.0	117	12	US-10-424-599-206883,	Sequence 206883,	168	19	95.0	193	14	US-10-156-761-13355	Sequence 13355, A
96	19	95.0	117	12	US-10-424-599-221584,	Sequence 221584,	169	19	95.0	193	15	US-10-377-078-38	Sequence 38, Appl
97	19	95.0	118	12	US-10-424-599-271082,	Sequence 271082,	170	19	95.0	195	12	US-10-424-599-231551	Sequence 231551,
98	19	95.0	119	12	US-10-424-599-240644,	Sequence 240644,	171	19	95.0	195	12	US-10-424-599-238939	Sequence 238939,
99	19	95.0	119	14	US-10-563-568-12	Sequence 12, Appl1	172	19	95.0	195	15	US-10-369-493-3234	Sequence 3234, Ap
100	19	95.0	120	12	US-10-424-599-189718,	Sequence 189718,	173	19	95.0	197	15	US-09-966-8808-2	Sequence 2, Appl1
101	19	95.0	120	12	US-10-424-599-276785,	Sequence 276785,	174	19	95.0	198	12	US-10-425-114-47425	Sequence 47425, A
102	19	95.0	122	11	US-09-864-408A-526	Sequence 526, App	175	19	95.0	199	12	US-10-405-877-85	Sequence 85, Appl
103	19	95.0	122	11	US-10-424-599-173522	Sequence 173522,	176	19	95.0	200	12	US-10-424-599-176742	Sequence 176742,
104	19	95.0	127	12	US-10-424-599-178767	Sequence 178767,	177	19	95.0	200	12	US-10-424-599-220990,	Sequence 220990,
105	19	95.0	129	15	US-10-299-636-68	Sequence 68, Appl	178	19	95.0	201	12	US-10-276-162-2	Sequence 2, Appl1
106	19	95.0	130	12	US-10-424-599-184879	Sequence 184879,	179	19	95.0	201	15	US-10-085-198-134	Sequence 134, App
107	19	95.0	131	9	US-09-738-626-5204	Sequence 193911,	180	19	95.0	201	15	US-10-282-122A-60863	Sequence 60863, A
108	19	95.0	131	9	US-09-738-626-5204	Sequence 5204, Ap	181	19	95.0	202	12	US-10-424-599-231852	Sequence 231852,
109	19	95.0	133	9	US-09-864-761-33708	Sequence 932, App	182	19	95.0	204	12	US-10-424-599-231852	Sequence 231852,
110	19	95.0	133	9	US-10-424-599-158331	Sequence 158331,	183	19	95.0	204	12	US-10-424-599-285941	Sequence 285941,
111	19	95.0	134	9	US-09-815-242-10513	Sequence 10513, A	184	19	95.0	204	12	US-10-634-548-31	Sequence 31, Appl
112	19	95.0	135	12	US-10-424-599-271153	Sequence 271153,	185	19	95.0	204	12	US-10-634-548-31	Sequence 31, Appl
113	19	95.0	135	12	US-10-425-114-53378	Sequence 53378, A	186	19	95.0	205	9	US-09-922-217-1108	Sequence 1108, Ap
114	19	95.0	138	12	US-10-112-944-912	Sequence 912, App	187	19	95.0	205	12	US-10-296-115-1330	Sequence 1330, Ap
115	19	95.0	139	14	US-10-156-761-8842	Sequence 8842, Ap	188	19	95.0	205	12	US-10-748-748	Sequence 748, Ap
116	19	95.0	139	15	US-10-385-415-120	Sequence 120, App	189	19	95.0	205	12	US-10-335-977-7980	Sequence 7980, Ap
117	19	95.0	140	15	US-10-108-260A-3390	Sequence 3390, Ap	190	19	95.0	205	12	US-10-335-977-7980	Sequence 1108, Ap
118	19	95.0	143	15	US-10-385-415-119	Sequence 119, App	191	19	95.0	208	12	US-10-282-122A-44800	Sequence 44800, A
119	19	95.0	145	13	US-10-013-379-11	Sequence 11, Appl	192	19	95.0	208	14	US-10-156-761-135621	Sequence 135621, A
120	19	95.0	145	15	US-10-289-762-167	Sequence 167, App	193	19	95.0	209	12	US-10-282-122A-72561	Sequence 72561, A
121	19	95.0	146	12	US-10-282-122A-70260	Sequence 70260, A	194	19	95.0	209	12	US-10-424-599-267539	Sequence 267539,
122	19	95.0	146	15	US-10-139-794-87	Sequence 87, Appl	195	19	95.0	212	14	US-10-127-032-130	Sequence 130, App
123	19	95.0	148	12	US-10-424-599-143742	Sequence 143742,	196	19	95.0	213	9	US-09-943-671-17	Sequence 17, Appl
124	19	95.0	148	12	US-10-424-599-219032	Sequence 219032,	197	19	95.0	213	9	US-09-738-626-3673	Sequence 3673, Ap
125	19	95.0	152	12	US-10-424-599-283372	Sequence 283372,	198	19	95.0	213	14	US-10-043-487-364	Sequence 364, App
126	19	95.0	152	13	US-10-016-157A-151	Sequence 151, App	199	19	95.0	213	14	US-10-405-877-17	Sequence 17, Appl
127	19	95.0	156	12	US-10-424-599-254161	Sequence 254161,	200	19	95.0	215	14	US-10-106-698-5887	Sequence 5887, Ap
128	19	95.0	157	14	US-10-238-075-357	Sequence 357, App	201	19	95.0	215	14	US-09-738-626-5708	Sequence 5708, Ap
129	19	95.0	159	12	US-10-424-599-256669	Sequence 256669,	202	19	95.0	215	12	US-10-282-122A-62976	Sequence 62976, A
130	19	95.0	163	12	US-10-424-599-169320	Sequence 169320,	203	19	95.0	215	12	US-10-326-671-69	Sequence 69, Appl
131	19	95.0	164	9	US-09-738-973-203	Sequence 203, App	204	19	95.0	215	14	US-10-424-599-162041	Sequence 162041,
132	19	95.0	164	9	US-09-854-133-203	Sequence 203, App	205	19	95.0	220	12	US-10-384-060-76	Sequence 76, Appl
133	19	95.0	164	11	US-09-864-408A-1148	Sequence 1148, Ap	206	19	95.0	221	14	US-10-424-599-266598	Sequence 266598,
134	19	95.0	164	14	US-10-144-649A-203	Sequence 203, App	207	19	95.0	221	14	US-10-156-761-12750	Sequence 12750, A
135	19	95.0	165	12	US-10-424-599-148999	Sequence 148999,	208	19	95.0	221	14	US-10-424-599-266598	Sequence 1539, Ap
136	19	95.0	165	12	US-10-424-599-234773	Sequence 234773,	209	19	95.0	221	14	US-10-156-761-12750	Sequence 157122,
137	19	95.0	167	9	US-09-925-300-981	Sequence 981, App	210	19	95.0	222	9	US-09-925-300-1639	Sequence 36974, A
138	19	95.0	168	12	US-10-424-599-198953	Sequence 198953,	211	19	95.0	222	12	US-10-424-599-157122	Sequence 52543, A
139	19	95.0	172	14	US-10-425-114-57575	Sequence 57575, A	212	19	95.0	222	12	US-10-425-114-36974	Sequence 52543, A
140	19	95.0	172	14	US-10-101-482-16	Sequence 16, Appl	213	19	95.0	222	12	US-10-425-114-32549	Sequence 4621, Ap
141	19	95.0	173	12	US-10-424-599-158029	Sequence 158029,	214	19	95.0	224	12	US-10-424-599-151189	Sequence 151189,
142	19	95.0	175	12	US-10-335-977-5868	Sequence 5868, Ap	215	19	95.0	224	15	US-10-094-749-1944	Sequence 1944, Ap
143	19	95.0	175	14	US-10-156-761-12575	Sequence 12575, A	216	19	95.0	224	15	US-10-351-334-155	Sequence 262650,
144	19	95.0	175	14	US-10-247-671-184	Sequence 184, App	217	19	95.0	225	12	US-10-351-334-155	Sequence 155, App
145	19	95.0	175	14	US-09-864-761-33919	Sequence 33919, A	218	19	95.0	225	12	US-10-424-599-262650	Sequence 14916, A
146	19	95.0	178	9	US-09-925-301-914	Sequence 914, App	219	19	95.0	225	12	US-10-156-761-14916	Sequence 11674, A
147	19	95.0	178	9	US-09-925-301-914	Sequence 592, App	220	19	95.0	225	15	US-10-369-493-14989	Sequence 14289, A
148	19	95.0	178	15	US-10-264-049-3578	Sequence 3578, Ap	221	19	95.0	225	15	US-10-369-493-14989	Sequence 14859, A
149	19	95.0	180	12	US-10-424-599-170244	Sequence 170244,	222	19	95.0	226	9	US-09-738-626-4027	Sequence 4027, Ap
150	19	95.0	180	14	US-10-156-761-12228	Sequence 12228, A	223	19	95.0	226	9	US-10-425-114-5374	Sequence 53174, A
151	19	95.0	180	12	US-10-335-977-7979	Sequence 7979, Ap	224	19	95.0	226	9	US-10-425-114-5374	Sequence 53174, A
152	19	95.0	184	12	US-10-335-977-7979	Sequence 17, Appl	225	19	95.0	226	9	US-10-425-114-5374	Sequence 53174, A
153	19	95.0	187	9	US-09-804-156-17	Sequence 9, Appl1	226	19	95.0	226	9	US-10-425-114-5374	Sequence 53174, A
154	19	95.0	187	9	US-09-804-156-17	Sequence 9, Appl1	227	19	95.0	226	9	US-10-425-114-5374	Sequence 53174, A
155	19	95.0	187	13	US-10-125-459-9	Sequence 9, Appl1	228	19	95.0	227	16	US-10-425-114-47426	Sequence 47426, A
156	19	95.0	187	13	US-10-067-761-17	Sequence 17, Appl	229	19	95.0	227	16	US-10-425-114-47426	Sequence 47426, A
157	19	95.0	187	14	US-10-319-519-17	Sequence 17, Appl	230	19	95.0	228	12	US-10-051-325-8	Sequence 8, Appl1
158	19	95.0	188	9	US-09-734-569-134	Sequence 134, App	231	19	95.0	228	12	US-10-424-599-143683	Sequence 143683,
159	19	95.0	188	12	US-10-282-122A-76586	Sequence 76586, A	232	19	95.0	229	12	US-10-424-599-143683	Sequence 4027, Ap
160	19	95.0	188	12	US-10-282-122A-78177	Sequence 78177, A	233	19	95.0	230	9	US-09-738-626-4027	Sequence 53174, A
161	19	95.0	188	12	US-10-424-599-256741	Sequence 256741,	234	19	95.0	231	12	US-10-425-114-57318	Sequence 57318, A

235	19	95.0	232	10	US-09-232-088-40	Sequence 40, Appl	308	19	95.0	270	9	US-09-738-626-1943	Sequence 3943, Ap
236	19	95.0	232	15	US-10-340-792-40	Sequence 40, Appl	309	19	95.0	270	12	US-10-262-511-148	Sequence 148, App
237	19	95.0	234	12	US-10-425-114-45098	Sequence 45098, A	310	19	95.0	270	15	US-10-369-493-10566	Sequence 10566, A
238	19	95.0	234	12	US-10-425-114-63330	Sequence 63330, A	311	19	95.0	271	15	US-10-264-237-2658	Sequence 2658, Ap
239	19	95.0	234	12	US-10-425-114-64490	Sequence 64490, A	312	19	95.0	272	10	US-09-764-891-2960	Sequence 2960, Ap
240	19	95.0	236	9	US-09-815-242-11738	Sequence 11738, A	313	19	95.0	272	10	US-10-282-122A-70355	Sequence 70355, A
241	19	95.0	236	12	US-10-412-599B-11410	Sequence 11738, A	314	19	95.0	273	14	US-10-182-447-3	Sequence 3, Appl1
242	19	95.0	236	12	US-10-412-599B-11410	Sequence 1410, Ap	315	19	95.0	273	12	US-10-282-122A-51559	Sequence 51559, A
243	19	95.0	236	15	US-10-374-780A-1221	Sequence 308, App	316	19	95.0	274	12	US-10-282-122A-71371	Sequence 71371, A
244	19	95.0	238	12	US-10-282-122A-70662	Sequence 1221, Ap	317	19	95.0	275	12	US-10-282-122A-52045	Sequence 52045, A
245	19	95.0	238	12	US-10-282-122A-72399	Sequence 70662, A	318	19	95.0	275	12	US-10-424-599-146071	Sequence 146071, A
246	19	95.0	238	14	US-10-137-036-68	Sequence 72399, A	319	19	95.0	278	12	US-10-282-122A-65461	Sequence 65461, A
247	19	95.0	239	12	US-10-424-599-283468	Sequence 68, Appl	320	19	95.0	278	12	US-10-282-122A-65761	Sequence 65761, A
248	19	95.0	240	12	US-10-338-777-39	Sequence 283468, A	321	19	95.0	279	15	US-10-369-493-11888	Sequence 11888, A
249	19	95.0	241	12	US-10-282-122A-56234	Sequence 39, Appl	322	19	95.0	280	9	US-09-738-626-4505	Sequence 4505, Ap
250	19	95.0	241	12	US-10-425-114-36715	Sequence 56234, A	323	19	95.0	280	15	US-10-369-493-20769	Sequence 20769, A
251	19	95.0	242	12	US-10-425-114-47756	Sequence 36715, A	324	19	95.0	281	13	US-10-059-964-58	Sequence 58, Appl
252	19	95.0	243	10	US-09-870-406A-44	Sequence 47756, A	325	19	95.0	281	14	US-10-156-761-9102	Sequence 9102, Ap
253	19	95.0	243	12	US-10-424-599-257488	Sequence 44, Appl	326	19	95.0	282	14	US-09-285-300-1694	Sequence 1694, Ap
254	19	95.0	244	14	US-10-159-901-44	Sequence 44, Appl	327	19	95.0	282	12	US-10-282-122A-47275	Sequence 47275, A
255	19	95.0	244	12	US-10-282-122A-72187	Sequence 72187, A	328	19	95.0	282	12	US-10-156-761-12355	Sequence 12355, A
256	19	95.0	244	12	US-10-424-599-158126	Sequence 158126, A	329	19	95.0	282	8	US-08-781-986A-5203	Sequence 5203, Ap
257	19	95.0	246	12	US-09-771-161A-129	Sequence 271081, A	330	19	95.0	283	12	US-10-425-114-43988	Sequence 43988, A
258	19	95.0	246	9	US-09-838-312-9	Sequence 129, App	331	19	95.0	283	12	US-10-329-624-5203	Sequence 5203, Ap
259	19	95.0	246	14	US-10-156-761-13663	Sequence 13663, A	332	19	95.0	283	12	US-10-425-114-58685	Sequence 58685, A
260	19	95.0	246	15	US-10-421-138A-311	Sequence 311, App	333	19	95.0	284	12	US-10-282-122A-49530	Sequence 49530, A
261	19	95.0	246	15	US-10-374-780A-1244	Sequence 1244, Ap	334	19	95.0	285	12	US-10-282-122A-69922	Sequence 69922, A
262	19	95.0	248	12	US-10-425-114-44065	Sequence 44065, A	335	19	95.0	285	12	US-10-425-114-64235	Sequence 64235, A
263	19	95.0	249	9	US-09-838-312-9	Sequence 9, Appli	336	19	95.0	285	12	US-10-282-122A-69316	Sequence 69316, A
264	19	95.0	249	11	US-09-829-495-9	Sequence 9, Appli	337	19	95.0	286	12	US-10-362-511-144	Sequence 144, App
265	19	95.0	249	14	US-10-156-761-9204	Sequence 9, Appli	338	19	95.0	286	12	US-10-424-599-17161	Sequence 147161, A
266	19	95.0	249	15	US-10-409-701-15	Sequence 9204, Ap	339	19	95.0	287	15	US-10-369-493-20884	Sequence 20884, A
267	19	95.0	249	15	US-10-310-154-606	Sequence 15, Appl	340	19	95.0	288	12	US-10-282-122A-66912	Sequence 66912, A
268	19	95.0	253	9	US-09-815-242-10279	Sequence 606, App	341	19	95.0	288	12	US-10-425-114-66363	Sequence 66363, A
269	19	95.0	253	15	US-10-369-493-23537	Sequence 10279, A	342	19	95.0	289	9	US-09-815-242-11193	Sequence 11193, A
270	19	95.0	254	12	US-10-424-599-189853	Sequence 158554, A	343	19	95.0	289	12	US-10-282-122A-58429	Sequence 58429, A
271	19	95.0	254	12	US-10-262-511-150	Sequence 150, App	344	19	95.0	289	12	US-10-425-114-47186	Sequence 47186, A
272	19	95.0	254	12	US-10-369-493-11429	Sequence 11429, A	345	19	95.0	289	12	US-10-425-114-47186	Sequence 47186, A
273	19	95.0	255	15	US-10-096-623-40	Sequence 40, Appl	346	19	95.0	290	12	US-10-424-599-196081	Sequence 196081, A
274	19	95.0	255	14	US-10-260-539-40	Sequence 40, Appl	347	19	95.0	290	15	US-10-369-493-15865	Sequence 15865, A
275	19	95.0	256	15	US-10-421-138A-312	Sequence 19054, A	348	19	95.0	291	15	US-10-041-018-27	Sequence 27, Appl
276	19	95.0	256	15	US-10-421-138A-312	Sequence 312, App	349	19	95.0	291	14	US-10-017-161-2430	Sequence 2430, Ap
277	19	95.0	256	15	US-09-738-626-6642	Sequence 1247, Ap	350	19	95.0	291	15	US-10-292-798-2070	Sequence 2070, Ap
278	19	95.0	257	9	US-09-738-626-6642	Sequence 6642, Ap	351	19	95.0	291	15	US-10-564-237-2770	Sequence 2770, Ap
279	19	95.0	257	14	US-10-106-698-5221	Sequence 5221, Ap	352	19	95.0	292	15	US-10-369-493-5341	Sequence 5341, Ap
280	19	95.0	257	15	US-10-139-794-16	Sequence 16, Appl	353	19	95.0	292	15	US-10-369-493-15494	Sequence 15494, A
281	19	95.0	257	15	US-09-815-242-11889	Sequence 11889, A	354	19	95.0	293	12	US-10-282-122A-76466	Sequence 76466, A
282	19	95.0	258	12	US-10-282-122A-66426	Sequence 66426, A	355	19	95.0	293	14	US-10-156-761-14726	Sequence 14726, A
283	19	95.0	258	12	US-10-424-599-248369	Sequence 248369, A	356	19	95.0	294	12	US-10-399-867-20	Sequence 20, Appl
284	19	95.0	258	14	US-10-424-599-248369	Sequence 303, App	357	19	95.0	294	14	US-10-336-699-12	Sequence 12, Appl
285	19	95.0	260	12	US-10-097-111-303	Sequence 46566, A	358	19	95.0	294	14	US-10-306-762-155	Sequence 155, App
286	19	95.0	260	12	US-10-425-114-46566	Sequence 19, Appl	359	19	95.0	294	15	US-10-369-493-13821	Sequence 13821, A
287	19	95.0	261	10	US-09-893-519A-19	Sequence 69377, A	360	19	95.0	295	14	US-10-306-762-141	Sequence 141, App
288	19	95.0	261	12	US-10-282-122A-69377	Sequence 69377, A	361	19	95.0	295	14	US-10-282-122A-50243	Sequence 50243, A
289	19	95.0	261	14	US-10-011-364-10	Sequence 10, Appl	362	19	95.0	297	12	US-10-282-122A-60194	Sequence 60194, A
290	19	95.0	261	15	US-10-369-493-12238	Sequence 12238, A	363	19	95.0	297	12	US-10-282-122A-60194	Sequence 60194, A
291	19	95.0	261	15	US-10-321-962-10	Sequence 10, Appl	364	19	95.0	297	12	US-10-425-114-41342	Sequence 41342, A
292	19	95.0	262	12	US-09-800-729-88	Sequence 8, Appl	365	19	95.0	299	9	US-09-393-634-35	Sequence 35, Appl
293	19	95.0	263	10	US-09-746-783-2	Sequence 2, Appli	366	19	95.0	299	10	US-09-310-332-1	Sequence 1, Appli
294	19	95.0	263	11	US-09-833-245-2213	Sequence 2, Appli	367	19	95.0	299	10	US-09-832-522-73	Sequence 73, Appl
295	19	95.0	263	14	US-10-189-447-1	Sequence 1, Appli	368	19	95.0	299	12	US-10-364-861-35	Sequence 35, Appl
296	19	95.0	264	12	US-10-282-122A-52414	Sequence 52414, A	369	19	95.0	299	13	US-10-059-964-46	Sequence 46, Appl
297	19	95.0	265	12	US-10-425-114-41437	Sequence 41437, A	370	19	95.0	299	14	US-10-314-639-46	Sequence 46, Appl
298	19	95.0	265	12	US-10-282-122A-52414	Sequence 52414, A	371	19	95.0	299	14	US-10-383-982-35	Sequence 35, Appl
299	19	95.0	265	12	US-10-425-114-62005	Sequence 62005, A	372	19	95.0	303	9	US-09-738-626-3734	Sequence 3734, Ap
300	19	95.0	266	12	US-10-424-599-252581	Sequence 252581, A	373	19	95.0	303	12	US-10-425-114-53342	Sequence 53342, A
301	19	95.0	267	12	US-10-282-122A-78178	Sequence 78178, A	374	19	95.0	304	12	US-10-424-599-219921	Sequence 219921, A
302	19	95.0	267	12	US-10-425-114-45104	Sequence 45104, A	375	19	95.0	304	12	US-10-425-114-62081	Sequence 62081, A
303	19	95.0	267	12	US-10-421-138A-315	Sequence 66738, A	376	19	95.0	304	15	US-10-369-493-9667	Sequence 9667, Ap
304	19	95.0	267	15	US-10-421-138A-315	Sequence 315, App	377	19	95.0	305	9	US-09-738-626-6591	Sequence 6591, Ap
305	19	95.0	268	9	US-09-745-763-164	Sequence 164, App	378	19	95.0	306	12	US-10-425-114-62473	Sequence 62473, A
306	19	95.0	268	16	US-10-253-662-19	Sequence 19, Appl	379	19	95.0	306	12	US-10-425-114-64440	Sequence 64440, A
307	19	95.0	269	15	US-10-104-047-2182	Sequence 2182, Ap	380	19	95.0	306	12	US-10-425-114-67258	Sequence 67258, A

361	19	95.0	306	14	US-10-091-007-40	Sequence 40, Appl	454	19	95.0	321	14	US-10-025-806-32	Sequence 32, Appl
362	19	95.0	306	15	US-10-320-797-3084	Sequence 3084, Ap	455	19	95.0	322	14	US-10-282-122A-53145	Sequence 53145, A
363	19	95.0	307	14	US-10-156-761-8528	Sequence 8528, Ap	456	19	95.0	322	14	US-10-270-861-7	Sequence 7, Appl
364	19	95.0	308	9	US-09-935-299-930	Sequence 930, Appl	457	19	95.0	322	14	US-10-270-861-11	Sequence 11, Appl
365	19	95.0	308	9	US-09-738-626-6728	Sequence 6728, Ap	458	19	95.0	322	14	US-10-282-122A-69906	Sequence 69906, A
366	19	95.0	308	14	US-09-925-299-930	Sequence 930, Appl	459	19	95.0	323	12	US-10-424-599-238938	Sequence 238938, A
367	19	95.0	308	14	US-10-032-201B-253	Sequence 253, Appl	460	19	95.0	323	12	US-10-282-122A-57182	Sequence 57182, A
368	19	95.0	309	12	US-10-425-114-43522	Sequence 43522, A	461	19	95.0	324	12	US-10-282-122A-61308	Sequence 61308, A
369	19	95.0	309	12	US-10-023-601-96	Sequence 96, Appl	462	19	95.0	324	14	US-10-156-761-13399	Sequence 13399, A
370	19	95.0	309	15	US-10-369-493-7640	Sequence 7640, Ap	463	19	95.0	324	14	US-10-156-761-14083	Sequence 14083, A
371	19	95.0	309	15	US-10-369-493-11118	Sequence 11118, A	464	19	95.0	325	12	US-10-262-511-152	Sequence 152, Appl
372	19	95.0	310	9	US-09-836-909-2	Sequence 2, Appl	465	19	95.0	325	14	US-10-270-861-1	Sequence 1, Appl
373	19	95.0	310	12	US-10-282-122A-67405	Sequence 67405, A	466	19	95.0	325	14	US-10-270-861-13	Sequence 13, Appl
374	19	95.0	310	12	US-10-424-599-163497	Sequence 163497, A	467	19	95.0	326	15	US-10-369-493-1001	Sequence 1001, Ap
375	19	95.0	310	12	US-10-424-599-230916	Sequence 230916, A	468	19	95.0	327	12	US-10-424-599-214738	Sequence 214738, A
376	19	95.0	310	15	US-10-191-121-18	Sequence 18, Appl	469	19	95.0	327	12	US-10-425-114-52852	Sequence 52852, A
377	19	95.0	310	16	US-10-360-534-16	Sequence 16, Appl	470	19	95.0	327	14	US-10-311-657-8	Sequence 8, Appl
378	19	95.0	311	9	US-09-886-055-9	Sequence 9, Appl	471	19	95.0	328	9	US-09-815-242-10935	Sequence 10935, A
379	19	95.0	311	9	US-09-738-626-3671	Sequence 3671, Ap	472	19	95.0	328	10	US-09-765-061B-75	Sequence 75, Appl
380	19	95.0	311	10	US-09-804-291-9	Sequence 9, Appl	473	19	95.0	328	10	US-09-765-061B-76	Sequence 76, Appl
381	19	95.0	311	10	US-09-832-522-20	Sequence 20, Appl	474	19	95.0	328	12	US-10-282-122A-56878	Sequence 56878, A
382	19	95.0	311	10	US-09-795-271-25	Sequence 25, Appl	475	19	95.0	328	15	US-10-369-493-20838	Sequence 20838, A
383	19	95.0	311	11	US-09-981-151A-75	Sequence 75, Appl	476	19	95.0	328	15	US-10-369-493-20838	Sequence 20838, A
384	19	95.0	311	11	US-09-844-861A-12	Sequence 12, Appl	477	19	95.0	329	9	US-09-771-161A-128	Sequence 128, App
385	19	95.0	311	12	US-10-343-650A-364	Sequence 364, Appl	478	19	95.0	331	9	US-09-815-242-11761	Sequence 11761, A
386	19	95.0	311	12	US-10-343-650A-364	Sequence 364, Appl	479	19	95.0	331	12	US-10-424-599-176898	Sequence 176898, A
387	19	95.0	311	14	US-10-023-597-52	Sequence 52, Appl	480	19	95.0	331	12	US-10-425-114-43781	Sequence 43781, A
388	19	95.0	311	14	US-10-017-161-280	Sequence 280, Appl	481	19	95.0	334	15	US-10-264-049-2446	Sequence 2446, Ap
389	19	95.0	311	14	US-10-017-161-284	Sequence 284, Appl	482	19	95.0	335	10	US-09-984-271-118	Sequence 118, App
390	19	95.0	311	15	US-10-005-041A-20	Sequence 20, Appl	483	19	95.0	335	12	US-10-425-114-39152	Sequence 39152, A
391	19	95.0	311	15	US-10-292-798-250	Sequence 250, Appl	484	19	95.0	335	12	US-09-984-276-118	Sequence 118, App
392	19	95.0	311	15	US-10-292-798-254	Sequence 254, Appl	485	19	95.0	336	12	US-10-369-493-48792	Sequence 48792, A
393	19	95.0	312	9	US-09-836-055-187	Sequence 187, Appl	486	19	95.0	336	12	US-10-282-122A-67767	Sequence 67767, A
394	19	95.0	312	9	US-09-738-626-5007	Sequence 5007, Ap	487	19	95.0	336	12	US-10-282-122A-78253	Sequence 78253, A
395	19	95.0	312	10	US-09-804-291-187	Sequence 187, Appl	488	19	95.0	336	12	US-10-425-114-46555	Sequence 46555, A
396	19	95.0	312	12	US-10-282-122A-53997	Sequence 53997, A	489	19	95.0	336	15	US-10-369-493-419	Sequence 419, App
397	19	95.0	312	12	US-10-182-822A-17	Sequence 17, Appl	490	19	95.0	337	10	US-09-793-705-42	Sequence 42, Appl
398	19	95.0	312	12	US-10-231-079-34	Sequence 34, Appl	491	19	95.0	337	12	US-10-425-114-57499	Sequence 57499, A
399	19	95.0	312	12	US-10-343-650A-656	Sequence 656, Appl	492	19	95.0	337	12	US-10-425-114-59818	Sequence 59818, A
400	19	95.0	312	14	US-10-017-161-50	Sequence 50, Appl	493	19	95.0	337	12	US-10-425-114-63730	Sequence 63730, A
401	19	95.0	312	14	US-10-017-161-498	Sequence 498, Appl	494	19	95.0	337	15	US-10-369-493-11195	Sequence 11195, A
402	19	95.0	312	15	US-10-387-629-16	Sequence 16, Appl	495	19	95.0	338	9	US-09-852-976-1	Sequence 1, Appl
403	19	95.0	312	15	US-10-292-798-42	Sequence 42, Appl	496	19	95.0	338	12	US-10-425-114-53202	Sequence 53202, A
404	19	95.0	313	12	US-10-369-493-11089	Sequence 11089, A	497	19	95.0	338	12	US-10-425-114-61153	Sequence 61153, A
405	19	95.0	313	12	US-10-282-122A-45385	Sequence 45385, A	498	19	95.0	338	12	US-10-425-114-62750	Sequence 62750, A
406	19	95.0	314	12	US-10-424-599-181578	Sequence 181578, A	499	19	95.0	338	12	US-10-425-114-66618	Sequence 66618, A
407	19	95.0	314	12	US-10-424-599-185947	Sequence 185947, A	500	19	95.0	339	9	US-09-832-312-3	Sequence 3, Appl
408	19	95.0	315	14	US-10-424-599-185944	Sequence 185944, A	501	19	95.0	339	9	US-09-832-312-34	Sequence 34, Appl
409	19	95.0	315	14	US-10-023-597-102	Sequence 102, Appl	502	19	95.0	339	9	US-09-832-312-36	Sequence 36, Appl
410	19	95.0	315	14	US-10-238-075-812	Sequence 812, Appl	503	19	95.0	339	9	US-09-832-312-38	Sequence 38, Appl
411	19	95.0	315	14	US-10-311-626-1	Sequence 1, Appl	504	19	95.0	339	9	US-09-832-312-40	Sequence 40, Appl
412	19	95.0	315	16	US-10-309-290-138	Sequence 138, Appl	505	19	95.0	339	11	US-09-829-495-3	Sequence 3, Appl
413	19	95.0	316	12	US-10-072-012-407	Sequence 407, Appl	506	19	95.0	339	11	US-09-829-495-34	Sequence 34, Appl
414	19	95.0	316	14	US-10-236-699-16	Sequence 16, Appl	507	19	95.0	339	11	US-09-829-495-36	Sequence 36, Appl
415	19	95.0	317	12	US-10-425-114-43447	Sequence 43447, A	508	19	95.0	339	11	US-09-829-495-38	Sequence 38, Appl
416	19	95.0	318	10	US-09-795-271-40	Sequence 40, Appl	509	19	95.0	339	11	US-09-829-495-40	Sequence 40, Appl
417	19	95.0	318	10	US-09-007-218-63	Sequence 63, Appl	510	19	95.0	339	14	US-10-157-031-387	Sequence 387, Appl
418	19	95.0	318	11	US-09-965-422-41	Sequence 41, Appl	511	19	95.0	339	14	US-10-156-761-14680	Sequence 14680, A
419	19	95.0	318	11	US-09-981-566A-42	Sequence 42, Appl	512	19	95.0	339	14	US-10-053-288-34	Sequence 34, Appl
420	19	95.0	318	11	US-09-981-566A-40	Sequence 40, Appl	513	19	95.0	339	14	US-10-446-848-5	Sequence 5, Appl
421	19	95.0	318	14	US-10-032-201B-284	Sequence 284, Appl	514	19	95.0	342	12	US-10-425-114-48872	Sequence 48872, A
422	19	95.0	318	14	US-10-375-611-1	Sequence 1, Appl	515	19	95.0	342	12	US-10-276-774-2289	Sequence 2289, Ap
423	19	95.0	318	14	US-10-247-671-131	Sequence 131, Appl	516	19	95.0	342	12	US-09-805-020-49	Sequence 49, Appl
424	19	95.0	319	9	US-09-832-312-5	Sequence 5, Appl	517	19	95.0	342	14	US-10-238-075-285	Sequence 285, App
425	19	95.0	319	11	US-09-829-495-5	Sequence 5, Appl	518	19	95.0	345	9	US-09-925-300-1188	Sequence 1188, App
426	19	95.0	319	12	US-10-425-114-70828	Sequence 70828, A	519	19	95.0	345	12	US-10-424-599-23433	Sequence 23433, A
427	19	95.0	319	13	US-10-001-426-4	Sequence 4, Appl	520	19	95.0	346	10	US-09-953-348-107	Sequence 107, App
428	19	95.0	319	13	US-10-003-014-4	Sequence 4, Appl	521	19	95.0	346	12	US-10-425-114-65218	Sequence 65218, A
429	19	95.0	319	14	US-10-424-599-186648	Sequence 186648, A	522	19	95.0	346	12	US-10-425-114-67904	Sequence 67904, A
430	19	95.0	320	12	US-10-335-977-5833	Sequence 5833, Ap	523	19	95.0	346	14	US-10-267-255-107	Sequence 107, App
431	19	95.0	320	12	US-10-461-990-28	Sequence 28, Appl	524	19	95.0	347	9	US-09-738-626-6932	Sequence 6932, Ap
432	19	95.0	320	12	US-10-424-599-163498	Sequence 163498, A	525	19	95.0	347	12	US-10-425-114-52822	Sequence 52822, A
433	19	95.0	321	12			526	19	95.0	347	14	US-10-306-762-240	Sequence 240, App

527	19	95.0	347	15	US-10-264-049-2882	Sequence 2882, Ap	600	19	95.0	384	15	US-10-369-493-18459	Sequence 18459, A
528	19	95.0	349	12	US-10-335-977-5814	Sequence 5814, Ap	601	19	95.0	385	12	US-10-425-114-67317	Sequence 67317, A
529	19	95.0	349	16	US-10-389-566-1121	Sequence 1121, Ap	602	19	95.0	385	15	US-10-369-493-6719	Sequence 6719, A
530	19	95.0	351	16	US-10-389-566-2148	Sequence 2148, Ap	603	19	95.0	386	9	US-09-815-242-11041	Sequence 11041, A
531	19	95.0	349	12	US-10-282-122A-46888	Sequence 46888, A	604	19	95.0	386	12	US-10-282-122A-58201	Sequence 58201, A
532	19	95.0	351	15	US-10-365-493-10659	Sequence 10659, A	605	19	95.0	386	15	US-10-369-493-4047	Sequence 4047, Ap
533	19	95.0	351	15	US-10-369-493-10888	Sequence 10888, A	606	19	95.0	389	12	US-10-424-599-264035	Sequence 264035, A
534	19	95.0	352	12	US-10-282-122A-60673	Sequence 60673, A	607	19	95.0	389	12	US-10-425-114-59654	Sequence 59654, A
535	19	95.0	352	15	US-10-369-493-20240	Sequence 20240, A	608	19	95.0	389	14	US-10-166-225A-82	Sequence 82, Appl
536	19	95.0	352	16	US-10-389-566-2165	Sequence 2165, Ap	609	19	95.0	389	14	US-10-166-225A-83	Sequence 83, Appl
537	19	95.0	353	14	US-10-270-861-9	Sequence 9, Appl1	610	19	95.0	390	15	US-10-369-493-20787	Sequence 20787, Ap
538	19	95.0	354	16	US-10-615-624-4	Sequence 4, Appl1	611	19	95.0	390	15	US-10-369-493-8879	Sequence 8879, Ap
539	19	95.0	355	14	US-10-108-698-4846	Sequence 4846, Ap	612	19	95.0	390	15	US-10-369-493-18574	Sequence 18574, A
540	19	95.0	357	12	US-10-363-616-4348	Sequence 438, App	613	19	95.0	391	15	US-10-365-493-3160	Sequence 3160, Ap
541	19	95.0	358	12	US-10-087-152-1845	Sequence 1845, Ap	614	19	95.0	392	12	US-10-425-114-50194	Sequence 50194, A
542	19	95.0	359	14	US-10-032-585-7893	Sequence 7893, Ap	615	19	95.0	392	15	US-10-369-493-11600	Sequence 11600, A
543	19	95.0	360	12	US-10-425-114-65637	Sequence 65637, A	616	19	95.0	392	15	US-10-369-493-14301	Sequence 14301, A
544	19	95.0	363	12	US-10-425-114-61874	Sequence 61874, A	617	19	95.0	392	15	US-10-369-493-14467	Sequence 14467, A
545	19	95.0	364	12	US-10-424-599-214841	Sequence 214841, A	618	19	95.0	392	15	US-10-369-493-15069	Sequence 15069, A
546	19	95.0	364	13	US-10-086-623-38	Sequence 38, Appl1	619	19	95.0	392	15	US-10-369-493-15069	Sequence 15069, A
547	19	95.0	365	14	US-10-260-539-38	Sequence 38, Appl1	620	19	95.0	393	12	US-10-087-193-1923	Sequence 1923, Ap
548	19	95.0	365	14	US-10-156-761-8947	Sequence 8947, Ap	621	19	95.0	394	9	US-09-934-903-4	Sequence 4, Appl1
549	19	95.0	367	12	US-10-425-114-64842	Sequence 64842, A	622	19	95.0	394	9	US-09-934-868-64	Sequence 64, Appl1
550	19	95.0	368	12	US-10-424-599-227208	Sequence 227208, A	623	19	95.0	394	10	US-09-934-947A-8	Sequence 8, Appl1
551	19	95.0	368	15	US-10-369-493-145	Sequence 145, App	624	19	95.0	394	12	US-10-282-122A-54557	Sequence 54557, A
552	19	95.0	368	15	US-10-369-493-1745	Sequence 2742, Ap	625	19	95.0	394	12	US-10-425-114-50770	Sequence 4, Appl1
553	19	95.0	369	12	US-10-282-122A-56898	Sequence 56898, A	626	19	95.0	394	15	US-10-700-003-4	Sequence 103, App
554	19	95.0	369	12	US-10-282-122A-73956	Sequence 73956, A	627	19	95.0	394	15	US-10-381-779-103	Sequence 103, App
555	19	95.0	369	12	US-10-282-122A-74479	Sequence 74479, A	628	19	95.0	394	15	US-10-381-779-121	Sequence 121, App
556	19	95.0	369	14	US-10-115-571A-31	Sequence 31, Appl	629	19	95.0	394	15	US-10-264-049-2612	Sequence 2612, Ap
557	19	95.0	369	14	US-10-446-836-37	Sequence 37, Appl	630	19	95.0	394	15	US-10-369-493-8612	Sequence 8, Appl1
558	19	95.0	370	9	US-09-808-972-4	Sequence 4, Appl1	631	19	95.0	395	12	US-10-369-567-8	Sequence 236415
559	19	95.0	370	11	US-09-876-813-53	Sequence 53, Appl1	632	19	95.0	396	14	US-10-424-599-236415	Sequence 88, Appl1
560	19	95.0	370	12	US-10-608-095-4	Sequence 4, Appl1	633	19	95.0	396	14	US-10-204-887-88	Sequence 152, App
561	19	95.0	370	12	US-10-368-095-4	Sequence 4, Appl1	634	19	95.0	397	12	US-09-769-742A-152	Sequence 47722, A
562	19	95.0	370	13	US-10-086-623-36	Sequence 36, Appl1	635	19	95.0	397	12	US-10-282-122A-47722	Sequence 47722, A
563	19	95.0	370	13	US-10-039-847A-4	Sequence 4, Appl1	636	19	95.0	397	12	US-10-282-122A-74248	Sequence 74248, A
564	19	95.0	370	14	US-10-260-539-36	Sequence 36, Appl1	637	19	95.0	397	15	US-10-365-493-18573	Sequence 18573, A
565	19	95.0	370	14	US-10-274-659-4	Sequence 4, Appl1	638	19	95.0	398	14	US-10-166-225A-74	Sequence 74, Appl1
566	19	95.0	370	14	US-10-226-538-4	Sequence 4, Appl1	639	19	95.0	399	12	US-10-424-599-212357	Sequence 212357, A
567	19	95.0	370	14	US-10-011-364-8	Sequence 8, Appl1	640	19	95.0	401	12	US-10-087-152-1926	Sequence 152, App
568	19	95.0	370	15	US-10-321-962-8	Sequence 8, Appl1	641	19	95.0	401	12	US-10-424-599-225710	Sequence 225710, A
569	19	95.0	371	12	US-10-293-122A-72020	Sequence 72020, A	642	19	95.0	402	12	US-10-282-122A-55610	Sequence 55610, A
570	19	95.0	371	12	US-10-282-122A-73684	Sequence 73684, A	643	19	95.0	402	12	US-10-424-599-225710	Sequence 225710, A
571	19	95.0	371	12	US-10-282-122A-74303	Sequence 74303, A	644	19	95.0	402	12	US-10-282-122A-55610	Sequence 55610, A
572	19	95.0	371	12	US-10-424-599-218495	Sequence 218495, A	645	19	95.0	403	15	US-10-365-493-21706	Sequence 21706, A
573	19	95.0	372	9	US-09-800-729-213	Sequence 213, App	646	19	95.0	403	15	US-10-365-493-21706	Sequence 21706, A
574	19	95.0	372	10	US-09-765-061B-74	Sequence 74, Appl1	647	19	95.0	404	12	US-10-282-122A-68690	Sequence 68690, A
575	19	95.0	372	10	US-09-765-061B-78	Sequence 78, Appl1	648	19	95.0	404	12	US-10-282-122A-68690	Sequence 68690, A
576	19	95.0	374	9	US-09-815-242-13447	Sequence 13447, A	649	19	95.0	405	12	US-10-389-647-635	Sequence 13613, A
577	19	95.0	374	9	US-09-815-242-13692	Sequence 13692, A	650	19	95.0	405	12	US-10-389-647-635	Sequence 635, App
578	19	95.0	374	12	US-10-425-114-38772	Sequence 38772, A	651	19	95.0	405	15	US-10-405-877-63	Sequence 63, Appl1
579	19	95.0	374	14	US-10-329-386-33411	Sequence 33411, A	652	19	95.0	405	15	US-10-369-493-13889	Sequence 13889, A
580	19	95.0	375	9	US-09-815-242-10928	Sequence 10928, A	653	19	95.0	406	16	US-10-282-122A-57253	Sequence 57253, A
581	19	95.0	377	12	US-10-289-122A-53791	Sequence 53791, A	654	19	95.0	406	16	US-10-282-122A-57253	Sequence 436, App
582	19	95.0	378	9	US-09-371-536-46	Sequence 46, Appl1	655	19	95.0	406	16	US-10-287-226-438	Sequence 438, App
583	19	95.0	378	15	US-10-369-493-13902	Sequence 13902, A	656	19	95.0	407	14	US-10-287-226-438	Sequence 74832, A
584	19	95.0	379	12	US-10-282-122A-51871	Sequence 51871, A	657	19	95.0	407	12	US-10-282-122A-74832	Sequence 74832, A
585	19	95.0	379	12	US-10-282-122A-51871	Sequence 51871, A	658	19	95.0	407	12	US-10-424-599-193438	Sequence 193438, A
586	19	95.0	379	12	US-10-282-122A-77140	Sequence 77140, A	659	19	95.0	407	12	US-10-424-599-193438	Sequence 8848, Ap
587	19	95.0	379	12	US-10-205-331-68	Sequence 68, Appl1	660	19	95.0	408	14	US-10-156-761-8848	Sequence 8848, Ap
588	19	95.0	379	14	US-10-157-223-6	Sequence 6, Appl1	661	19	95.0	408	14	US-10-156-761-8848	Sequence 95, Appl1
589	19	95.0	380	9	US-09-815-242-11291	Sequence 11291, A	662	19	95.0	409	14	US-10-081-816-14	Sequence 95, Appl1
590	19	95.0	380	10	US-09-882-227-466	Sequence 466, App	663	19	95.0	409	16	US-10-617-217A-95	Sequence 5136, Ap
591	19	95.0	381	12	US-10-282-122A-61191	Sequence 61191, A	664	19	95.0	409	16	US-09-765-061B-74	Sequence 5136, Ap
592	19	95.0	381	12	US-10-425-114-65079	Sequence 65079, A	665	19	95.0	408	9	US-09-765-061B-74	Sequence 5136, Ap
593	19	95.0	382	12	US-10-282-122A-61883	Sequence 61883, A	666	19	95.0	408	14	US-10-156-761-816-14	Sequence 9342, Ap
594	19	95.0	382	15	US-10-369-493-8452	Sequence 8452, Ap	667	19	95.0	408	14	US-10-156-761-816-14	Sequence 9342, Ap
595	19	95.0	383	12	US-10-282-122A-43909	Sequence 43909, A	668	19	95.0	409	14	US-10-081-816-14	Sequence 9342, Ap
596	19	95.0	383	14	US-10-369-493-8452	Sequence 8452, Ap	669	19	95.0	409	14	US-10-081-816-14	Sequence 9342, Ap
597	19	95.0	383	16	US-10-166-225A-76	Sequence 76, Appl1	670	19	95.0	410	14	US-10-424-599-266089	Sequence 266089, A
598	19	95.0	384	16	US-10-389-566-1853	Sequence 1853, Ap	671	19	95.0	410	14	US-10-424-599-266089	Sequence 266089, A
599	19	95.0	384	16	US-10-389-566-1853	Sequence 1853, Ap	672	19	95.0	410	14	US-10-081-816-23	Sequence 23, Appl1
599	19	95.0	384	16	US-10-389-566-1853	Sequence 1853, Ap	672	19	95.0	410	14	US-10-081-816-23	Sequence 23, Appl1
599	19	95.0	384	16	US-10-389-566-1853	Sequence 1853, Ap	672	19	95.0	410	14	US-10-081-816-23	Sequence 23, Appl1
599	19	95.0	384	16	US-10-389-566-1853	Sequence 1853, Ap	672	19	95.0	410	14	US-10-081-816-23	Sequence 23, Appl1
599	19	95.0	384	16	US-10-389-566-1853	Sequence 1853, Ap	672	19	95.0	410	14	US-10-081-816-23	Sequence 23, Appl1
599	19	95.0	384	16	US-10-389-566-1853	Sequence 1853, Ap	672	19	95.0	410	14	US-10-081-816-23	Sequence 23, Appl1
599	19	95.0	384	16	US-10-389-566-1853	Sequence 1853, Ap	672	19	95.0	410	14	US-10-081-816-23	Sequence 23, Appl1
599	19	95.0	384	16	US-10-389-566-1853	Sequence 1853, Ap	672	19	95.0	410	14	US-10-081-816-23	Sequence 23, Appl1
599	19	95.0	384	16	US-10-389-566-1853	Sequence 1853, Ap	672	19	95.0	410	14	US-10-081-816-23	Sequence 23, Appl1
599	19	95.0	384	16	US-10-389-566-1853	Sequence 1853, Ap							

673	19	95.0	413	12	US-10-282-122A-48732	Sequence 48732, A	746	19	95.0	439	12	US-10-282-122A-55457	Sequence 55457, A
674	19	95.0	413	15	US-10-369-493-1140	Sequence 1140, Ap	747	19	95.0	439	12	US-10-282-122A-77845	Sequence 77845, A
675	19	95.0	414	9	US-09-820-893-69	Sequence 69, Appl	748	19	95.0	440	12	US-10-282-122A-77845	Sequence 77845, A
676	19	95.0	414	12	US-10-424-599-285027	Sequence 285027, A	749	19	95.0	441	12	US-10-424-599-161614	Sequence 161614, A
677	19	95.0	414	12	US-10-607-565-69	Sequence 69, Appl	750	19	95.0	441	12	US-10-424-599-161615	Sequence 161615, A
678	19	95.0	414	16	US-10-287-228-48	Sequence 440, App	751	19	95.0	441	12	US-10-424-599-226388	Sequence 226388, A
679	19	95.0	415	12	US-10-282-122A-55349	Sequence 55349, A	752	19	95.0	444	12	US-10-282-122A-56949	Sequence 56949, A
680	19	95.0	415	12	US-10-282-122A-56787	Sequence 56787, A	753	19	95.0	444	12	US-10-369-493-124	Sequence 124, App
681	19	95.0	415	12	US-10-282-122A-68742	Sequence 68742, A	754	19	95.0	445	11	US-09-981-151A-73	Sequence 73, Appl
682	19	95.0	416	14	US-10-156-761-12502	Sequence 12502, A	755	19	95.0	445	12	US-10-282-122A-45092	Sequence 45092, A
683	19	95.0	416	15	US-10-369-493-4737	Sequence 4737, Ap	756	19	95.0	445	12	US-10-282-122A-74539	Sequence 74539, A
684	19	95.0	416	15	US-10-369-493-4882	Sequence 4882, Ap	757	19	95.0	446	12	US-10-282-122A-48392	Sequence 48392, A
685	19	95.0	416	15	US-10-369-493-4736	Sequence 4736, Ap	758	19	95.0	446	12	US-10-282-122A-48392	Sequence 48392, A
686	19	95.0	417	12	US-10-424-599-274599	Sequence 274599, A	759	19	95.0	446	12	US-10-425-114-52177	Sequence 52177, A
687	19	95.0	417	12	US-10-425-114-56059	Sequence 56059, A	760	19	95.0	447	9	US-09-908-928-4	Sequence 4, Appl
688	19	95.0	417	14	US-10-128-714-3345	Sequence 3345, Ap	761	19	95.0	447	9	US-09-972-528-4	Sequence 4, Appl
689	19	95.0	417	14	US-10-128-714-8345	Sequence 8345, Ap	762	19	95.0	447	12	US-10-282-122A-72069	Sequence 72069, A
690	19	95.0	418	11	US-09-981-151A-29	Sequence 29, Appl	763	19	95.0	447	12	US-10-282-122A-72107	Sequence 72107, A
691	19	95.0	418	15	US-10-369-493-20475	Sequence 20475, A	764	19	95.0	447	12	US-10-164-966-17	Sequence 17, Appl
692	19	95.0	420	15	US-10-369-493-13561	Sequence 13561, A	765	19	95.0	447	14	US-10-156-761-12559	Sequence 12559, A
693	19	95.0	421	9	US-08-815-242-5303	Sequence 5303, Ap	766	19	95.0	448	12	US-10-282-122A-48951	Sequence 48951, A
694	19	95.0	421	9	US-08-815-242-12388	Sequence 12388, A	767	19	95.0	448	12	US-10-424-599-266964	Sequence 266964, A
695	19	95.0	421	9	US-08-815-242-12820	Sequence 12820, A	768	19	95.0	448	14	US-10-238-075-1420	Sequence 1420, Ap
696	19	95.0	421	12	US-10-282-122A-71223	Sequence 71223, A	769	19	95.0	449	12	US-10-425-114-38657	Sequence 38657, A
697	19	95.0	421	12	US-10-425-114-56431	Sequence 56431, A	770	19	95.0	450	15	US-10-369-493-2490	Sequence 2490, Ap
698	19	95.0	421	14	US-10-138-701-8	Sequence 8, Appl	771	19	95.0	453	9	US-09-738-626-5635	Sequence 5635, Ap
699	19	95.0	421	14	US-10-138-701-22	Sequence 22, Appl	772	19	95.0	453	12	US-10-282-122A-72044	Sequence 72044, A
700	19	95.0	422	12	US-10-282-122A-66702	Sequence 66702, A	773	19	95.0	453	12	US-10-282-122A-74651	Sequence 74651, A
701	19	95.0	423	12	US-10-282-122A-53283	Sequence 53283, A	774	19	95.0	454	12	US-10-425-114-63654	Sequence 63654, A
702	19	95.0	423	12	US-10-424-599-204268	Sequence 204268, A	775	19	95.0	454	12	US-09-981-151A-72	Sequence 72, Appl
703	19	95.0	423	12	US-10-425-114-46740	Sequence 46740, A	776	19	95.0	455	11	US-10-282-122A-55920	Sequence 55920, A
704	19	95.0	423	12	US-10-107-431-49	Sequence 49, Appl	777	19	95.0	456	12	US-10-282-122A-78329	Sequence 78329, A
705	19	95.0	423	12	US-10-183-687-222	Sequence 222, App	778	19	95.0	456	15	US-10-320-797-3204	Sequence 3204, Ap
706	19	95.0	423	15	US-10-369-493-5557	Sequence 5557, Ap	779	19	95.0	457	9	US-09-888-615-110	Sequence 110, App
707	19	95.0	424	12	US-10-282-122A-45014	Sequence 45014, A	780	19	95.0	457	11	US-09-981-151A-71	Sequence 71, Appl
708	19	95.0	424	12	US-10-424-599-214840	Sequence 214840, A	781	19	95.0	457	12	US-10-282-122A-67190	Sequence 67190, A
709	19	95.0	424	14	US-10-156-761-14870	Sequence 14870, A	782	19	95.0	457	14	US-10-156-761-11073	Sequence 11073, A
710	19	95.0	425	14	US-10-156-761-17776	Sequence 17776, Ap	783	19	95.0	457	14	US-10-156-761-13584	Sequence 13584, A
711	19	95.0	425	14	US-10-214-446-20	Sequence 20, Appl	784	19	95.0	458	15	US-10-369-493-8104	Sequence 8104, Ap
712	19	95.0	426	12	US-10-425-114-48506	Sequence 48506, A	785	19	95.0	459	12	US-10-282-122A-45167	Sequence 45167, Ap
713	19	95.0	427	15	US-10-369-493-29777	Sequence 2977, Ap	786	19	95.0	460	12	US-10-425-114-56969	Sequence 56969, A
714	19	95.0	427	15	US-10-369-493-11890	Sequence 11890, A	787	19	95.0	460	12	US-10-369-493-8005	Sequence 8005, Ap
715	19	95.0	427	15	US-10-104-047-2262	Sequence 11890, A	788	19	95.0	460	15	US-10-108-260A-2540	Sequence 2540, Ap
716	19	95.0	428	12	US-10-403-571-1110	Sequence 2262, Ap	789	19	95.0	461	12	US-10-282-122A-50440	Sequence 50440, A
717	19	95.0	430	15	US-10-369-493-17335	Sequence 110, App	790	19	95.0	461	12	US-10-259-194A-186	Sequence 186, App
718	19	95.0	431	12	US-10-282-122A-47233	Sequence 17335, A	791	19	95.0	461	15	US-10-424-599-266524	Sequence 266524, A
719	19	95.0	431	12	US-10-282-122A-51920	Sequence 47233, A	792	19	95.0	465	14	US-10-195-141-19	Sequence 19, Appl
720	19	95.0	431	12	US-10-282-122A-51920	Sequence 51920, A	793	19	95.0	465	15	US-10-345-072-19	Sequence 61, Appl
721	19	95.0	431	12	US-10-444-599-212351	Sequence 212351, A	794	19	95.0	465	12	US-10-282-122A-63275	Sequence 61, Appl
722	19	95.0	431	12	US-10-041-018-260	Sequence 260, App	795	19	95.0	468	15	US-10-120-801-61	Sequence 1416, Ap
723	19	95.0	431	15	US-10-369-493-21530	Sequence 21530, A	796	19	95.0	468	15	US-10-369-493-1416	Sequence 1416, Ap
724	19	95.0	432	12	US-10-425-114-72646	Sequence 72646, A	797	19	95.0	469	12	US-10-417-700A-55	Sequence 65, Appl
725	19	95.0	433	15	US-10-369-493-8573	Sequence 8573, Ap	798	19	95.0	469	12	US-10-417-700A-55	Sequence 77, Appl
726	19	95.0	433	15	US-10-138-714-8545	Sequence 8545, Ap	799	19	95.0	469	14	US-10-032-585-7098	Sequence 7098, Ap
727	19	95.0	433	15	US-10-369-493-10203	Sequence 10203, A	800	19	95.0	471	12	US-10-282-122A-76782	Sequence 76782, A
728	19	95.0	434	12	US-10-282-122A-66904	Sequence 66904, A	801	19	95.0	471	12	US-10-425-114-72794	Sequence 72794, A
729	19	95.0	435	12	US-10-627-476-238	Sequence 238, App	802	19	95.0	472	11	US-09-981-151A-27	Sequence 27, Appl
730	19	95.0	436	12	US-10-282-122A-50291	Sequence 50291, A	803	19	95.0	472	12	US-10-424-599-251597	Sequence 251597, A
731	19	95.0	436	12	US-10-425-114-46821	Sequence 46821, A	804	19	95.0	472	12	US-10-425-114-62714	Sequence 62714, A
732	19	95.0	436	12	US-10-041-018-235	Sequence 235, App	805	19	95.0	472	15	US-10-421-654-68	Sequence 68, Appl
733	19	95.0	436	15	US-10-369-493-21378	Sequence 21378, A	806	19	95.0	474	12	US-10-282-122A-62083	Sequence 62083, A
734	19	95.0	437	9	US-09-918-932-2	Sequence 2, Appl	807	19	95.0	475	12	US-10-425-114-40744	Sequence 40744, A
735	19	95.0	437	9	US-09-738-626-4226	Sequence 4226, Ap	808	19	95.0	475	12	US-10-425-114-39831	Sequence 39831, A
736	19	95.0	437	12	US-10-282-122A-42940	Sequence 42940, A	809	19	95.0	475	12	US-10-424-599-257731	Sequence 257731, A
737	19	95.0	437	12	US-10-282-122A-73421	Sequence 73421, A	810	19	95.0	477	14	US-10-156-761-14307	Sequence 14307, A
738	19	95.0	437	12	US-10-450-055-6	Sequence 6, Appl	811	19	95.0	478	12	US-10-425-114-38072	Sequence 38072, A
739	19	95.0	437	14	US-10-177-293-457	Sequence 457, App	812	19	95.0	480	9	US-09-820-893-108	Sequence 108, App
740	19	95.0	437	15	US-10-296-027-1180	Sequence 1180, App	813	19	95.0	480	12	US-10-424-599-166523	Sequence 166523, A
741	19	95.0	438	12	US-10-282-122A-56153	Sequence 56153, A	814	19	95.0	480	12	US-10-425-114-63937	Sequence 63937, A
742	19	95.0	438	12	US-10-282-122A-56161	Sequence 56161, A	815	19	95.0	480	12	US-10-425-114-63937	Sequence 108, App
743	19	95.0	438	12	US-10-282-122A-75202	Sequence 75202, A	816	19	95.0	481	12	US-10-424-599-166520	Sequence 166520, A
744	19	95.0	438	12	US-10-282-122A-75609	Sequence 75609, A	817	19	95.0	481	12	US-10-282-122A-48208	Sequence 48208, A
745	19	95.0	439	11	US-09-764-875-930	Sequence 930, App	818	19	95.0	484	12	US-10-425-114-42223	Sequence 42223, A



819	19	95.0	485	9	US-09-769-864-1	Sequence 1, Appli	892	19	95.0	518	12	US-10-424-599-172544	Sequence 172544, A
820	19	95.0	485	9	US-09-769-864-2	Sequence 2, Appli	893	19	95.0	522	15	US-10-099-322-124	Sequence 124, App
821	19	95.0	485	9	US-09-769-864-6	Sequence 6, Appli	894	19	95.0	522	15	US-10-099-322-125	Sequence 125, App
822	19	95.0	485	9	US-09-769-864-7	Sequence 7, Appli	895	19	95.0	522	15	US-10-044-564-124	Sequence 124, App
823	19	95.0	485	9	US-09-769-864-8	Sequence 8, Appli	896	19	95.0	522	15	US-10-044-564-125	Sequence 125, App
824	19	95.0	485	9	US-09-854-346-2	Sequence 2, Appli	897	19	95.0	523	15	US-10-099-322-122	Sequence 122, App
825	19	95.0	485	9	US-09-854-346-4	Sequence 4, Appli	898	19	95.0	523	15	US-10-099-322-123	Sequence 123, App
826	19	95.0	485	9	US-09-854-346-12	Sequence 12, Appli	899	19	95.0	523	15	US-10-044-564-122	Sequence 122, App
827	19	95.0	485	9	US-09-854-346-13	Sequence 13, Appli	900	19	95.0	523	15	US-10-044-564-123	Sequence 123, App
828	19	95.0	485	9	US-09-902-188A-1	Sequence 1, Appli	901	19	95.0	524	12	US-10-425-114-62567	Sequence 62567, A
829	19	95.0	485	9	US-09-902-188A-2	Sequence 2, Appli	902	19	95.0	524	12	US-10-425-114-62567	Sequence 62567, A
830	19	95.0	485	9	US-09-902-188A-7	Sequence 7, Appli	903	19	95.0	524	12	US-10-425-114-68132	Sequence 68132, A
831	19	95.0	485	9	US-09-918-543-1	Sequence 1, Appli	904	19	95.0	524	12	US-10-403-161-34	Sequence 34, Appli
832	19	95.0	485	9	US-09-918-543-2	Sequence 2, Appli	905	19	95.0	524	12	US-10-403-161-40	Sequence 40, Appli
833	19	95.0	485	9	US-09-918-543-12	Sequence 12, Appli	906	19	95.0	524	15	US-10-099-322-34	Sequence 34, Appli
834	19	95.0	485	9	US-09-795-211-1	Sequence 1, Appli	907	19	95.0	524	15	US-10-099-322-121	Sequence 121, App
835	19	95.0	485	9	US-09-795-211-2	Sequence 2, Appli	908	19	95.0	524	15	US-10-044-564-34	Sequence 34, Appli
836	19	95.0	485	10	US-09-925-576C-2	Sequence 2, Appli	909	19	95.0	524	15	US-10-044-564-121	Sequence 121, App
837	19	95.0	485	10	US-09-925-576C-4	Sequence 4, Appli	910	19	95.0	526	9	US-09-815-242-13696	Sequence 13696, A
838	19	95.0	485	10	US-09-925-576C-12	Sequence 12, Appli	911	19	95.0	526	12	US-10-282-122A-73677	Sequence 73677, A
839	19	95.0	485	10	US-09-925-576C-13	Sequence 13, Appli	912	19	95.0	526	12	US-10-282-122A-73677	Sequence 73677, A
840	19	95.0	485	12	US-10-665-667-1	Sequence 1, Appli	913	19	95.0	527	12	US-10-425-114-46419	Sequence 46419, A
841	19	95.0	485	12	US-10-665-667-2	Sequence 2, Appli	914	19	95.0	527	12	US-10-425-114-43548	Sequence 43548, A
842	19	95.0	485	12	US-10-665-667-5	Sequence 5, Appli	915	19	95.0	529	12	US-10-403-161-36	Sequence 36, Appli
843	19	95.0	485	12	US-10-665-667-7	Sequence 7, Appli	916	19	95.0	529	12	US-10-403-161-38	Sequence 38, Appli
844	19	95.0	485	12	US-10-665-667-8	Sequence 8, Appli	917	19	95.0	530	15	US-10-369-493-19139	Sequence 19139, A
845	19	95.0	485	12	US-10-025-648-1	Sequence 1, Appli	918	19	95.0	531	12	US-10-425-114-45814	Sequence 45814, A
846	19	95.0	485	12	US-10-025-648-2	Sequence 2, Appli	919	19	95.0	533	12	US-10-282-122A-61253	Sequence

Sequence 75373, A  
 Sequence 76352, A  
 Sequence 65931, A  
 Sequence 21529, A  
 Sequence 1969, A  
 Sequence 34, Appl  
 Sequence 142, Appl  
 Sequence 19, Appl  
 Sequence 45750, A  
 Sequence 65302, A  
 Sequence 671, Appl  
 Sequence 1205, Appl  
 Sequence 40721, A  
 Sequence 42248, A  
 Sequence 186118  
 Sequence 8, Appl  
 Sequence 60, Appl  
 Sequence 7018, Appl  
 Sequence 11305, A  
 Sequence 58657, A  
 Sequence 7019, Appl  
 Sequence 598, Appl  
 Sequence 54815, A  
 Sequence 332, Appl  
 Sequence 25268, A  
 Sequence 4380, Appl  
 Sequence 2, Appl  
 Sequence 4, Appl  
 Sequence 6, Appl  
 Sequence 198064, A  
 Sequence 3, Appl  
 Sequence 10, Appl  
 Sequence 65911, A  
 Sequence 214104, A  
 Sequence 5869, Appl  
 Sequence 44744, A

## ALIGNMENTS

RESULT 1  
 ; Sequence 343, Application US/09791378  
 ; Patent No. US20020142303A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Parekh, Rajesh  
 ; TITLE OF INVENTION: PROTEINS, GENES AND THEIR USE FOR DIAGNOSIS AND TREATMENT OF SCHIZOPHRENIA  
 ; FILE REFERENCE: 9195-061-999  
 ; CURRENT APPLICATION NUMBER: US/09/791,378  
 ; CURRENT FILING DATE: 2001-02-23  
 ; PRIOR APPLICATION NUMBER: 09/750,395  
 ; PRIOR FILING DATE: 2000-12-28  
 ; NUMBER OF SEQ ID NOS: 677  
 ; SOFTWARE: Patentin version 3.0  
 ; SEQ ID NO 343  
 ; LENGTH: 8  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-791-378-343

Query Match 95.0%; Score 19; DB 9; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 Db 2 VAEF 5

RESULT 2  
 US-09-791-393-206

; Sequence 206, Application US/09791393  
 ; Publication No. US20030032200A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Herath, Herath Mudiyanseelage Athula Chandrasiri  
 ; APPLICANT: Parekh, Rajesh Bhikhu  
 ; APPLICANT: Rohlf, Christian  
 ; TITLE OF INVENTION: Proteins, Genes and Their Use for  
 ; TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)  
 ; TITLE OF INVENTION: and Unipolar Depression  
 ; FILE REFERENCE: 2543-1-001 N1  
 ; CURRENT APPLICATION NUMBER: US/09/791,393  
 ; CURRENT FILING DATE: 2002-01-02  
 ; EARLIER APPLICATION NUMBER: GB 0004412.3  
 ; EARLIER FILING DATE: 2000-02-24  
 ; EARLIER APPLICATION NUMBER: GB 0030050.9  
 ; EARLIER FILING DATE: 2000-12-08  
 ; EARLIER APPLICATION NUMBER: US 60/254,830  
 ; EARLIER FILING DATE: 2000-12-12  
 ; NUMBER OF SEQ ID NOS: 308  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 206  
 ; LENGTH: 8  
 ; TYPE: PRT  
 ; ORGANISM: homo sapien  
 ; US-09-791-393-206

Query Match 95.0%; Score 19; DB 10; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 Db 2 VAEF 5

## RESULT 3

US-09-791-389-206  
 ; Sequence 206, Application US/09791389  
 ; Publication No. US20030032773A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Herath, Herath Mudiyanseelage Athula Chandrasiri  
 ; APPLICANT: Parekh, Rajesh Bhikhu  
 ; APPLICANT: Rohlf, Christian  
 ; APPLICANT: Terrett, Jonathan Alexander  
 ; APPLICANT: Tyson, Kerry Louise  
 ; TITLE OF INVENTION: Proteins, Genes and Their Use for  
 ; TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)  
 ; TITLE OF INVENTION: and Unipolar Depression  
 ; FILE REFERENCE: 2543-1-001 N2  
 ; CURRENT APPLICATION NUMBER: US/09/791,389  
 ; CURRENT FILING DATE: 2001-02-23  
 ; PRIOR APPLICATION NUMBER: GB 0004412.3  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: GB 0030050.9  
 ; PRIOR FILING DATE: 2000-12-08  
 ; PRIOR APPLICATION NUMBER: US 60/254,830  
 ; PRIOR FILING DATE: 2000-12-12  
 ; NUMBER OF SEQ ID NOS: 308  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 206  
 ; LENGTH: 8  
 ; TYPE: PRT  
 ; ORGANISM: homo sapien  
 ; US-09-791-389-206

Query Match 95.0%; Score 19; DB 10; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 Db 2 VAEF 5

RESULT 4

S-10-128-711-86  
Sequence 86, Application US/10128711  
Publication No. US20030099634A1  
GENERAL INFORMATION:  
APPLICANT: VITIELLO, Maria A.

CHESTNUT, Robert W.  
SETTE, Alessandro D.  
CELIS, Esteban  
GRAY, Howard

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING CTL IMMUNITY

NUMBER OF SEQUENCES: 153

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Stuart Street Tower, One Market Plaza

CITY: San Francisco

STATE: California

COUNTRY: US

ZIP: 94105-1493

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/128,711

FILING DATE: 22-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/197,484

FILING DATE: 16-FEB-1994

APPLICATION NUMBER: US 07/935,811

FILING DATE: 26-AUG-1992

APPLICATION NUMBER: US 07/874,491

FILING DATE: 27-APR-1992

APPLICATION NUMBER: US 07/827,682

FILING DATE: 29-JAN-1992

APPLICATION NUMBER: US 07/749,568

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 14137-26-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 467-9600

TELEFAX: (206) 623-6793

INFORMATION FOR SEQ ID NO: 86:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 86:

S-10-128-711-86

Query Match 95.0%; Score 19; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5

b 2 VAEF 5

RESULT 5

S-10-128-711-145

Sequence 145, Application US/10128711

Publication No. US20030099634A1

GENERAL INFORMATION:

APPLICANT: VITIELLO, Maria A.

CHESTNUT, Robert W.  
SETTE, Alessandro D.  
CELIS, Esteban  
GRAY, Howard

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING CTL IMMUNITY

NUMBER OF SEQUENCES: 153

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourie and Crew

STREET: Stuart Street Tower, One Market Plaza

CITY: San Francisco

STATE: California

COUNTRY: US

ZIP: 94105-1493

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/128,711

FILING DATE: 22-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/197,484

FILING DATE: 16-FEB-1994

APPLICATION NUMBER: US 07/935,811

FILING DATE: 26-AUG-1992

APPLICATION NUMBER: US 07/874,491

FILING DATE: 27-APR-1992

APPLICATION NUMBER: US 07/827,682

FILING DATE: 29-JAN-1992

APPLICATION NUMBER: US 07/749,568

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 14137-26-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 467-9600

TELEFAX: (206) 623-6793

INFORMATION FOR SEQ ID NO: 145:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 145:

US-10-128-711-145

Query Match 95.0%; Score 19; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

DB 2 VAEF 5

RESULT 6

US-09-908-943A-197

Sequence 197, Application US/09908943A

Publication No. US20030017991A1

GENERAL INFORMATION:

APPLICANT: Yan, Riqiang

APPLICANT: Tomasselli, Alfredo G.

APPLICANT: Gurney, Mark E.

APPLICANT: Emmons, Thomas L.

APPLICANT: Bienkowski, Mike J.

APPLICANT: Heinrichson, Robert L.

TITLE OF INVENTION: SUBSTRATES AND ASSAYS FOR BETA-SECRETASE ACTIVITY

FILE REFERENCE: 29915/00281A.US1

```

; CURRENT APPLICATION NUMBER: US/09/908,943A
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/219,795
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 197
; LENGTH: 10
; TYPE: PRT
; ORGANISM: synthetic peptide sequence
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)..(4)
; OTHER INFORMATION: amino acid at position 4 has been derivatized with a statine
; NAME/KEY: SITE
; LOCATION: (10)..(10)
; OTHER INFORMATION: amino acid at position 10 has been derivatized with Bodipy FL
JS-09-908-943A-197

Query Match          95.0%; Score 19; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 5 VAEF 8

RESULT 7
US-09-908-943A-196
; Sequence 196, Application US/09908943A
; Publication No. US20030017991A1
; GENERAL INFORMATION:
; APPLICANT: Yan, Riciang
; APPLICANT: Tomasselli, Alfredo G.
; APPLICANT: Guiney, Mark E.
; APPLICANT: Emmons, Thomas L.
; APPLICANT: Bienkowski, Mike J.
; APPLICANT: Heinrichson, Robert L.
; TITLE OF INVENTION: SUBSTRATES AND ASSAYS FOR BETA-SECRETASE ACTIVITY
; FILE REFERENCE: 29915/00281A.US1
; CURRENT APPLICATION NUMBER: US/09/908,943A
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/219,795
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 196
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence: synthetic peptide sequence
; NAME/KEY: SITE
; LOCATION: (4)..(4)
; OTHER INFORMATION: amino acid at position 4 has been derivatized with a statine
US-09-908-943A-196

Query Match          95.0%; Score 19; DB 12; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 5 VAEF 8

RESULT 8
US-10-032-818-34
; Sequence 34, Application US/10032818
; Publication No. US20030092629A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.

```

```

; APPLICANT: Koelsch, Gerald
; APPLICANT: Ghosh, Arun K.
; TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof
; FILE REFERENCE: 2932.1006-007
; CURRENT APPLICATION NUMBER: US/10/032,818
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/275,756
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/258,705
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-032-818-34

Query Match          95.0%; Score 19; DB 14; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 5 VAEF 8

RESULT 9
US-10-239-313A-379
; Sequence 379, Application US/10239313A
; Publication No. US20030175285A1
; GENERAL INFORMATION:
; APPLICANT: KLINGUER - HAMOUR, Christine
; APPLICANT: BECK, Alain
; APPLICANT: CORVAIA, Nathalie
; APPLICANT: GOETSCH, Liliane
; TITLE OF INVENTION: MOLECULE OF PHARMACEUTICAL INTEREST COMPRISING AT ITS
; TITLE OF INVENTION: N-TERMINAL A GLUTAMIC ACID OR A GLUTAMINE IN THE FORM
; FILE REFERENCE: 343 727 - US
; CURRENT APPLICATION NUMBER: US/10/239,313A
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: FR 00/03711
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT 01/70772
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 697
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 379
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-239-313A-379

Query Match          95.0%; Score 19; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 4 VAEF 7

RESULT 10
US-10-119-528-7
; Sequence 7, Application US/10119528
; Publication No. US20030175722A1
; GENERAL INFORMATION:
; APPLICANT: Mann, M.
; APPLICANT: Mortensen, P.
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR SEARCHING GENOMIC DATABASES

```

FILE REFERENCE: MDSF-P01-004  
CURRENT APPLICATION NUMBER: US/10/119,528  
CURRENT FILING DATE: 2002-04-09  
PRIOR APPLICATION NUMBER: 60/282,551  
PRIOR FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: 60/285,362  
PRIOR FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 7  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
-10-119-528-7

Query Match	95.0%;	Score 19;	DB 14;	Length 15;
Best Local Similarity	100.0%;	Pred. No. 2e+02;		
Matches	4;	Conservative	0;	Mismatches 0;
		Indels	0;	Gaps 0;

2 VAEF 5  
10 VAEF 13

RESULT 11  
 -09-171-432A-62  
 Sequence 62, Application US/09171432A  
 Publication No. US20030181784A1  
 GENERAL INFORMATION:  
 APPLICANT: Fields, Howard A.  
 APPLICANT: Khudyakov, Yuri E.  
 TITLE OF INVENTION: Antigenically Reactive Regions of the  
 TITLE OF INVENTION: Hepatitis A Virus Polypeptide  
 NUMBER OF SEQUENCES: 88  
 CORRESPONDENCE ADDRESS:

NAME: Greene, Jamie L.  
REGISTRATION NUMBER: 32,457  
REFERENCE/DOCKET NUMBER: 03063-0233105  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 949-2400  
TELEFAX: (404) 949-2499  
INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /label= YK-1368  
09-171-432A-62

```

Query Match          95.0%; Score 19; DB 10; Length 20;
Best Local Similarity 100.0%; Pred.No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels

QY      2 VAEF 5
         ||||
Db      13 VAEF 16

```

RESULT 12  
 US-09-171-432A-63  
 ; Sequence 63, Application US/09171432A  
 ; Publication NO. US20030187184A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fields, Howard A.  
 ; APPLICANT: Khudiyakov, Yury E.  
 ; TITLE OF INVENTION: Antigenically Reactive Regions of the  
 ; TITLE OF INVENTION: Hepatitis A Virus Polyprotein  
 ; NUMBER OF SEQUENCES: 88  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Kilpatrick Stockton LLP  
 ; STREET: 3424 Peachtree Road, N.E.  
 ; CITY: Atlanta  
 ; STATE: Georgia  
 ; COUNTRY: USA  
 ; ZIP: 30326  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/171,432A  
 ; FILING DATE: 23-NOV-1998  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Greene, Jamie L.  
 ; REGISTRATION NUMBER: 32,467  
 ; REFERENCE/DOCKET NUMBER: 03063-0231US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (404) 949-2400  
 ; TELEFAX: (404) 949-2499  
 ; INFORMATION FOR SEQ ID NO: 63:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 20 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; FRAGMENT TYPE: internal  
 ; FEATURE:  
 ; NAME/KEY: Peptide  
 ; LOCATION: 1..20  
 ; OTHER INFORMATION: /label= YK-1369  
 US-09-171-432A-63

```

Query Match          95.0%; Score 19; DB 10; Length 20;
Best Local Similarity 100.0%; Pred.No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels
QY      2 VAEF 5
        ||||
Db      4 VAEF 7

```

RESULT 13  
US-09-171-432A-65  
; Sequence 65, Application US/09171432A  
; Publication No. US2003018718A1  
; GENERAL INFORMATION:  
; APPLICANT: Fields, Howard A.

APPLICANT: Khudyakov, Yury E.  
TITLE OF INVENTION: Hepatitically Reactive Regions of the  
TITLE OF INVENTION: Hepatitis A Virus Polypeptide  
NUMBER OF SEQUENCES: 88  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick Stockton LLP  
STREET: 3424 Peachtree Road, N.E.  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30326  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/171,432A  
FILING DATE: 23-NOV-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Greene, Jamie L.  
REGISTRATION NUMBER: 32,467  
REFERENCE/DOCKET NUMBER: 03063-0231US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 949-2400  
TELEFAX: (404) 949-2499  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..25  
OTHER INFORMATION: /label=YK-1832

JS-09-171-432A-65  
Query Match 95.0%; Score 19; DB 10; Length 25;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 3 VAEF 6

RESULT 14  
US-09-864-761-47201  
; Sequence 47201, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Ches, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 47201  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC005772.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.46  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.67  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.61  
US-09-864-761-47201

Query Match 95.0%; Score 19; DB 9; Length 28;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 4 VAEF 7

RESULT 15  
US-10-424-599-216297  
; Sequence 216297, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 216297  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_37345C.1.pap  
US-10-424-599-216297

Query Match 95.0%; Score 19; DB 12; Length 28;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02; Indels 0;  
Matches 4; Conservative 0; Mismatches 0; Gaps 0;

2 VAEF 5  
|||  
10 VAEF 13

US-10-242-355-570  
Sequence 69, Application US/10043344  
Publication No. US2003008086A1  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena M.  
APPLICANT: Harkness, Robin E.  
APPLICANT: Schryvers, Anthony B.  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Murdin, Andrew D.  
APPLICANT: Klein, Michel H.  
TITLE OF INVENTION: TRANSFERRIN RECEPTOR GENES  
FILE REFERENCE: 1038-1221 MIS  
CURRENT APPLICATION NUMBER: US/10/043,344  
CURRENT FILING DATE: 2002-07-01  
PRIOR APPLICATION NUMBER: 08/649,518  
PRIOR FILING DATE: 1996-05-17  
NUMBER OF SEQ ID NOS: 160  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 69  
LENGTH: 35  
TYPE: PRT  
ORGANISM: Haemophilus influenzae  
US-10-043-344-69

Query Match 95.0%; Score 19; DB 14; Length 35;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02; Indels 0;  
Matches 4; Conservative 0; Mismatches 0; Gaps 0;

2 VAEF 5  
|||  
14 VAEF 17

US-10-242-355-570  
Sequence 570, Application US/10242355  
Publication No. US2003023583A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC003C1  
CURRENT APPLICATION NUMBER: US/10/242,355  
CURRENT FILING DATE: 2002-09-13  
PRIOR APPLICATION NUMBER: 09/764,897  
PRIOR FILING DATE: 2001-01-17  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/214,886  
PRIOR FILING DATE: 2000-06-28  
PRIOR APPLICATION NUMBER: 60/217,487  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,758  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/220,963  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: 60/217,496  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,447  
PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/218,290  
PRIOR FILING DATE: 2000-07-14  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1267  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 570  
LENGTH: 37  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-242-355-570

Query Match 95.0%; Score 19; DB 15; Length 37;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02; Indels 0;  
Matches 4; Conservative 0; Mismatches 0; Gaps 0;

2 VAEF 5  
|||  
14 VAEF 17

US-10-430-752A-49  
Sequence 49, Application US/10430752A  
Publication No. US20040005669A1  
GENERAL INFORMATION:  
APPLICANT: Stahl, Stefan  
APPLICANT: Jonasson, Per  
APPLICANT: Nygren, Per-Ake  
APPLICANT: Uhlen, Mathias  
TITLE OF INVENTION: RECOMBINANT EXPRESSION OF INSULIN  
FILE REFERENCE: 11541-003001  
CURRENT APPLICATION NUMBER: US/10/430,752A  
CURRENT FILING DATE: 2003-05-06  
PRIOR APPLICATION NUMBER: 09/485,286  
PRIOR FILING DATE: 2000-02-07  
PRIOR APPLICATION NUMBER: PCT/GB98/02382  
PRIOR FILING DATE: 1998-08-07  
PRIOR APPLICATION NUMBER: GB 9716790.2  
PRIOR FILING DATE: 1997-08-07  
NUMBER OF SEQ ID NOS: 63  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 49  
LENGTH: 37  
TYPE: PRT  
ORGANISM: Verasper moseri  
US-10-430-752A-49

Query Match 95.0%; Score 19; DB 15; Length 37;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02; Indels 0;  
Matches 4; Conservative 0; Mismatches 0; Gaps 0;

2 VAEF 5  
|||  
24 VAEF 27

US-10-430-752A-45  
Sequence 45, Application US/10430752A  
Publication No. US20040005669A1  
GENERAL INFORMATION:  
APPLICANT: Stahl, Stefan  
APPLICANT: Jonasson, Per  
APPLICANT: Nygren, Per-Ake  
APPLICANT: Uhlen, Mathias  
TITLE OF INVENTION: RECOMBINANT EXPRESSION OF INSULIN  
FILE REFERENCE: 11541-003001  
CURRENT APPLICATION NUMBER: US/10/430,752A  
CURRENT FILING DATE: 2003-05-06  
PRIOR APPLICATION NUMBER: 09/485,286  
PRIOR FILING DATE: 2000-02-07

PRIOR APPLICATION NUMBER: PCT/GB98/02382  
PRIOR FILING DATE: 1998-08-07  
PRIOR APPLICATION NUMBER: GB 9716790.2  
PRIOR FILING DATE: 1997-08-07  
NUMBER OF SEQ ID NOS: 63  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 45  
LENGTH: 38  
TYPE: PRT  
ORGANISM: Lophius piscatorius  
IS-10-430-752A-45

Query Match 95.0%; Score 19; DB 15; Length 38;  
Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
|  
|  
|  
|  
b 25 VAEF 28

RESULT 20  
IS-10-372-003A-71  
Sequence 71, Application US/10372003A  
Publication No. US20030215846A1  
GENERAL INFORMATION:  
APPLICANT: Watt, Paul  
APPLICANT: Thomas, Wayne  
APPLICANT: Hopkins, Richard  
TITLE OF INVENTION: Methods of constructing and screening  
TITLE OF INVENTION: diverse expression libraries  
FILE REFERENCE: FBRI40.001CPI  
CURRENT APPLICATION NUMBER: US/10/372.003A  
CURRENT FILING DATE: 2003-02-21  
PRIOR APPLICATION NUMBER: US 09/568,229  
PRIOR FILING DATE: 2000-05-05  
PRIOR APPLICATION NUMBER: US 60/132,711  
PRIOR FILING DATE: 1999-05-05  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 71  
LENGTH: 41  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic clone BGf13 with first open reading  
OTHER INFORMATION: frame

Query Match 95.0%; Score 19; DB 15; Length 41;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
|  
|  
|  
|  
b 28 VAEF 31

RESULT 21  
IS-09-864-761-37202  
Sequence 37202, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aemica-x-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 37202  
LENGTH: 43  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL035530.11  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.86  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.95  
OTHER INFORMATION: SWISSPROT HIT: P33485, EVALUE 3.60e-00  
OTHER INFORMATION: EST\_HUMAN HIT: AI765888.1, EVALUE 3.00e-18  
US-09-864-761-37202

Query Match 95.0%; Score 19; DB 9; Length 43;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
|  
|  
|  
|  
b 40 VAEF 43

RESULT 22  
US-10-424-599-183032  
Sequence 183032, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B



CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 183032  
LENGTH: 43  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_136291C.1.pap  
1-10-424-599-183032

Query Match 95.0%; Score 19; DB 12; Length 43;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||  
25 VAEF 28

RESULT 23  
3-10-424-599-171318  
Sequence 171318, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 171318  
LENGTH: 45  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_125715C.1.pap  
3-10-424-599-171318

Query Match 95.0%; Score 19; DB 12; Length 45;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||  
20 VAEF 23

RESULT 24  
3-09-864-761-38166  
Sequence 38166, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
SEQ ID NO 38166  
LENGTH: 48  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL049833.3  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1  
OTHER INFORMATION: SWISSPROT HIT: P11298, EVALUATE 7.90e+00  
US-09-864-761-38166

Query Match 95.0%; Score 19; DB 9; Length 48;  
Best Local Similarity 100.0%; Pred. No. 6.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||  
3 VAEF 6

RESULT 25  
US-10-424-599-205632  
Sequence 205632, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 205632  
LENGTH: 52

```

TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_27713C.1.pep
US-10-424-599-205632

Query Match
Best Local Similarity 95.0%; Score 19; DB 12; Length 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 26 VAEF 29

RESULT 26
US-10-424-599-192247
; Sequence 192247, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 192247
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_1561C.1.pep
US-10-424-599-192247

Query Match
Best Local Similarity 95.0%; Score 19; DB 12; Length 55;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 50 VAEF 53

RESULT 27
US-10-424-599-233502
; Sequence 233502, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 233502
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_52879C.1.pep
US-10-424-599-233502

Query Match
Best Local Similarity 95.0%; Score 19; DB 12; Length 55;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 VAEF 5
DB 24 VAEF 27

RESULT 28
US-10-372-003A-63
; Sequence 63, Application US/10372003A
; Publication No. US20030215846A1
; GENERAL INFORMATION:
; APPLICANT: Watt, Paul
; APPLICANT: Thomas, Wayne
; APPLICANT: Hopkins, Richard
; TITLE OF INVENTION: Methods of constructing and screening
; TITLE OF INVENTION: diverse expression libraries
; FILE REFERENCE: FERIC40.001CPI
; CURRENT APPLICATION NUMBER: US/10/372,003A
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/568,229
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/132,711
; PRIOR FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic clone BGF05 with first open reading
; OTHER INFORMATION: frame
US-10-372-003A-63

Query Match
Best Local Similarity 95.0%; Score 19; DB 15; Length 55;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 28 VAEF 31

RESULT 29
US-10-372-003A-67
; Sequence 67, Application US/10372003A
; Publication No. US20030215846A1
; GENERAL INFORMATION:
; APPLICANT: Watt, Paul
; APPLICANT: Thomas, Wayne
; APPLICANT: Hopkins, Richard
; TITLE OF INVENTION: Methods of constructing and screening
; TITLE OF INVENTION: diverse expression libraries
; FILE REFERENCE: FERIC40.001CPI
; CURRENT APPLICATION NUMBER: US/10/372,003A
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/568,229
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/132,711
; PRIOR FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic clone BGF06 with first open reading
; OTHER INFORMATION: frame
US-10-372-003A-67

Query Match
Best Local Similarity 95.0%; Score 19; DB 15; Length 55;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; 2 VAEF 5
; ||||
; 28 VAEF 31
;
RESULT 30
; 10-219-329-4
; Sequence 4, Application US/10219329
; Publication No. US20030096757A1
; GENERAL INFORMATION:
; APPLICANT: Quirk, Stephen
; APPLICANT: Weart, Ilona f
; TITLE OF INVENTION: Anti-Cancer and Wound Healing Compounds
; FILE REFERENCE: 1443.035W01
; CURRENT APPLICATION NUMBER: US/10/219.329
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 10/032,376
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 60/312,726
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
;
; 3-10-219-329-4
;
; Query Match 95.0%; Score 19; DB 14; Length 56;
; Best Local Similarity 100.0%; Pred. No. 7.5e+02;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; 2 VAEF 5
; ||||
; 31 VAEF 34
;
;
; 3-10-153-185-4
; Sequence 4, Application US/10153185
; Publication No. US2003014859A1
; GENERAL INFORMATION:
; APPLICANT: Quirk, Stephen
; APPLICANT: Malik, Sohail
; TITLE OF INVENTION: Anti-Aging and Wound Healing Compounds
; FILE REFERENCE: 1443.034U81
; CURRENT APPLICATION NUMBER: US/10/153,185
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US 10/032,376
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 60/312,726
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
;
; 3-10-153-185-4
;
; Query Match 95.0%; Score 19; DB 14; Length 56;
; Best Local Similarity 100.0%; Pred. No. 7.5e+02;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; 2 VAEF 5
; ||||
; 31 VAEF 34
;
;
; 3-10-153-185-4
; Sequence 4, Application US/10153185
; Publication No. US2003014859A1
; GENERAL INFORMATION:
; APPLICANT: Quirk, Stephen
; APPLICANT: Malik, Sohail
; TITLE OF INVENTION: Anti-Aging and Wound Healing Compounds
; FILE REFERENCE: 1443.034U81
; CURRENT APPLICATION NUMBER: US/10/153,185
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US 10/032,376
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 60/312,726
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
;
; 3-10-153-185-4
;
; Query Match 95.0%; Score 19; DB 14; Length 56;
; Best Local Similarity 100.0%; Pred. No. 7.5e+02;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; 2 VAEF 5
; ||||
; 31 VAEF 34
;
;
; 3-10-219-561-4
; Sequence 4, Application US/10219561
; Publication No. US20030166567A1
; GENERAL INFORMATION:
; APPLICANT: Quirk, Stephen
; APPLICANT: Malik, Sohail
; APPLICANT: Villanueva, Julie M.
; TITLE OF INVENTION: Anti-Aging and Wound Healing Compounds
; FILE REFERENCE: 1443.008U52
; CURRENT APPLICATION NUMBER: US/10/219,561
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 10/032,376
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 10/153,185
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/312,726
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
;
; US-10-219-561-4
;
; Query Match 95.0%; Score 19; DB 14; Length 56;
; Best Local Similarity 100.0%; Pred. No. 7.5e+02;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 2 VAEF 5
; ||||
; DB 31 VAEF 34
;
;
; RESULT 33
; US-10-424-599-215734
; Sequence 215734, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 215734
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_36836C.1.pep
;
; US-10-424-599-215734
;
; Query Match 95.0%; Score 19; DB 12; Length 57;
; Best Local Similarity 100.0%; Pred. No. 7.6e+02;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 2 VAEF 5
; ||||
; DB 48 VAEF 51
;
;
; RESULT 34
; US-10-424-599-260762
; Sequence 260762, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
```

```
; Sequence 4, Application US/10219561
; Publication No. US20030166567A1
; GENERAL INFORMATION:
; APPLICANT: Quirk, Stephen
; APPLICANT: Malik, Sohail
; APPLICANT: Villanueva, Julie M.
; TITLE OF INVENTION: Anti-Aging and Wound Healing Compounds
; FILE REFERENCE: 1443.008U52
; CURRENT APPLICATION NUMBER: US/10/219,561
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 10/032,376
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 10/153,185
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/312,726
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
;
; US-10-219-561-4
;
; Query Match 95.0%; Score 19; DB 14; Length 56;
; Best Local Similarity 100.0%; Pred. No. 7.5e+02;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 2 VAEF 5
; ||||
; DB 31 VAEF 34
;
;
; RESULT 33
; US-10-424-599-215734
; Sequence 215734, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 215734
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_36836C.1.pep
;
; US-10-424-599-215734
;
; Query Match 95.0%; Score 19; DB 12; Length 57;
; Best Local Similarity 100.0%; Pred. No. 7.6e+02;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 2 VAEF 5
; ||||
; DB 48 VAEF 51
;
;
; RESULT 34
; US-10-424-599-260762
; Sequence 260762, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
```

APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 260762  
LENGTH: 57  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_77490C.1.pep  
US-10-424-599-260762

Query Match 95.0%; Score 19; DB 12; Length 57;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 4 VAEF 7

## RESULT 35

US-10-372-003A-75  
Sequence 75, Application US/10372003A

Publication No. US20030215846A1

GENERAL INFORMATION:

APPLICANT: Watt, Paul

APPLICANT: Thomas, Wayne

APPLICANT: Hopkins, Richard

TITLE OF INVENTION: Methods of constructing and screening

FILE OF INVENTION: diverse expression libraries

FILE REFERENCE: FBIC40.001CP1

CURRENT APPLICATION NUMBER: US/10/372,003A

CURRENT FILING DATE: 2003-02-21

PRIOR APPLICATION NUMBER: US 09/568,229

PRIOR FILING DATE: 2000-05-05

PRIOR APPLICATION NUMBER: US 60/132,711

PRIOR FILING DATE: 1999-05-05

NUMBER OF SEQ ID NOS: 81

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 75

LENGTH: 57

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: synthetic clone BGF24 with first open reading

OTHER INFORMATION: frame

US-10-372-003A-75

Query Match 95.0%; Score 19; DB 15; Length 57;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 28 VAEF 31

## RESULT 36

US-10-424-599-214930

Sequence 214930, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J

APPLICANT: Kovalic, David K

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 214930  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_36107C.1.pep  
US-10-424-599-214930

Query Match 95.0%; Score 19; DB 12; Length 58;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 13 VAEF 16

## RESULT 37

US-09-864-761-45291

Sequence 45291, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aecmica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 45291



; CURRENT FILING DATE: 2003-03-07  
; PRIOR APPLICATION NUMBER: US/09/393,634  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 92  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 80  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human GR24  
US-10-383-982-80

Query Match 95.0%; Score 19; DB 14; Length 68;  
Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 14 VAEF 17

RESULT 42  
US-09-864-761-45716  
; Sequence 45716, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Cherp, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 45716  
; LENGTH: 69  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL162171.1  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.81  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.91  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.85  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.72  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.71  
; OTHER INFORMATION: EST HUMAN HIT: BE55987.1, EVALUATE 2.00e-03  
; OTHER INFORMATION: SWISSPROT HIT: P56093, EVALUATE 2.00e-00  
US-09-864-761-45716

Query Match 95.0%; Score 19; DB 9; Length 69;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 15 VAEF 18

RESULT 43  
US-09-738-626-5596  
; Sequence 5596, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENO, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OKAZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 5596  
; LENGTH: 69  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-5596

Query Match 95.0%; Score 19; DB 9; Length 69;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 23 VAEF 26

RESULT 44  
US-10-424-599-185385  
; Sequence 185385, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 185385  
LENGTH: 69  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_138417C.1.pap  
US-10-424-599-185385

Query Match 95.0%; Score 19; DB 12; Length 69;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
Db 24 VAEF 27

RESULT 45  
US-10-029-386-28645  
Sequence 28645, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 28645  
LENGTH: 70  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO CHR17.1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.8  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.3  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.8  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.7  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.7  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8.7  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7  
OTHER INFORMATION: SWISSPROT HIT: P78716, EVALUE 3.00e-04  
US-10-029-386-28645

Query Match 95.0%; Score 19; DB 14; Length 70;  
Best Local Similarity 100.0%; Pred. No. 9.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
Db 43 VAEF 46

RESULT 46  
US-10-424-599-202605  
Sequence 202605, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 202605  
LENGTH: 71  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_24978C.1.pap  
US-10-424-599-202605

Query Match 95.0%; Score 19; DB 12; Length 71;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
Db 17 VAEF 20

RESULT 47  
US-10-424-599-274981  
Sequence 274981, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 274981  
LENGTH: 71  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(71)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_90329C.1.pap  
US-10-424-599-274981

Query Match 95.0%; Score 19; DB 12; Length 71;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
Db 39 VAEF 42

RESULT 48  
US-10-029-386-30348  
Sequence 30348, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: AEOMICA-X-2

CURRENT APPLICATION NUMBER: US/10/029,386

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 34288

SOFTWARE: Annonax Sequence Listing Engine vers. 1.1

SEQ ID NO 30348

LENGTH: 71

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO CHR7.1

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.82

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1

OTHER INFORMATION: EXPRESSED IN HUHT, SIGNAL = 0.73

OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 6.2

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1

OTHER INFORMATION: SWISSPROT HIT: P21333, EVALUATION 5.00e-22

US-10-029-386-30348

Query Match 95.0%; Score 19; DB 14; Length 71;

Best Local Similarity 100.0%; Pred. No. 9.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

DB 60 VAEF 63

RESULT 49

US-10-029-386-29320

Sequence 29320, Application US/10029386

Publication No. US2003019470A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: AECOMICA-X-2

CURRENT APPLICATION NUMBER: US/10/029,386

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 34288

SOFTWARE: Annonax Sequence Listing Engine vers. 1.1

SEQ ID NO 29320

LENGTH: 72

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO CHR2.1

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.98

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.99

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.6

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.81

OTHER INFORMATION: SWISSPROT HIT: P14650, EVALUATION 6.00e-14

US-10-029-386-29320

Query Match

Best Local Similarity 95.0%; Score 19; DB 14; Length 72;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

DB 45 VAEF 48

RESULT 50

US-10-214-188-10

Sequence 10, Application US/10214188

Publication No. US2003002260A1

GENERAL INFORMATION:

APPLICANT: LA THANGUE, NICHOLAS B.

BERNARDS, RENE  
HIOMANS, ELEANORE M.  
TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/10/214,188  
APPLICATION NUMBER: US/10/214,188  
FILING DATE: 08-Aug-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,139  
FILING DATE: 13-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 620-22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 74 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: Peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-10-214-188-10

Query Match 95.0%; Score 19; DB 14; Length 74;

Best Local Similarity 100.0%; Pred. No. 9.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

DB 23 VAEF 26

RESULT 51

US-10-156-761-9071

Sequence 9071, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 9071

LENGTH: 74

TYPE: PRT



ORGANISM: Streptomyces avermitilis  
3-10-156-761-9071

Query Match 95.0%; Score 19; DB 14; Length 74;  
Best Local Similarity 100.0%; Pred. No. 9.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
36 VAEF 39

RESULT 52  
US-10-424-599-156128  
Sequence 156128, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 156128  
LENGTH: 75  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(75)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_112004C.1.pep  
US-10-424-599-156128

Query Match 95.0%; Score 19; DB 12; Length 75;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
2 VAEF 5

RESULT 53  
3-10-424-599-261424  
Sequence 261424, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 261424  
LENGTH: 75  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_78089C.1.pep  
US-10-424-599-261424

Query Match 95.0%; Score 19; DB 12; Length 75;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
2 VAEF 5

RESULT 54  
US-10-424-599-253204  
Sequence 253204, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 253204  
LENGTH: 76  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(76)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_70669C.1.pep  
US-10-424-599-253204

Query Match 95.0%; Score 19; DB 12; Length 76;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
64 VAEF 67

RESULT 55  
US-10-424-599-211908  
Sequence 211908, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 211908  
LENGTH: 77  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_3337C.1.pep  
US-10-424-599-211908

Query Match 95.0%; Score 19; DB 12; Length 77;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
16 VAEF 19

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
12 VAEF 15

RESULT 54  
US-10-424-599-253204  
Sequence 253204, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 253204  
LENGTH: 76  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(76)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_70669C.1.pep  
US-10-424-599-253204

Query Match 95.0%; Score 19; DB 12; Length 76;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
64 VAEF 67

RESULT 55  
US-10-424-599-211908  
Sequence 211908, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 211908  
LENGTH: 77  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_3337C.1.pep  
US-10-424-599-211908

Query Match 95.0%; Score 19; DB 12; Length 77;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
16 VAEF 19

```
RESULT 56
US-10-335-977-7493
; Sequence 7493, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LANIWE & COCKFIELD
; STREET: 28 State Street
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 7493:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3) LOCATION 1...77
; SEQUENCE DESCRIPTION: SEQ ID NO: 7493:
US-10-335-977-7493
Query Match 95.0%; Score 19; DB 12; Length 77;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 40 VAEF 43

RESULT 57
US-10-029-386-32709
; Sequence 32709, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386

Query Match 95.0%; Score 19; DB 12; Length 77;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 40 VAEF 43

RESULT 58
US-10-424-599-162039
; Sequence 162039, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 162039
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_117339C.1.pcp
US-10-424-599-162039
Query Match 95.0%; Score 19; DB 12; Length 78;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 28 VAEF 31

RESULT 59
US-10-424-599-262701
; Sequence 262701, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 262701
; LENGTH: 78
```

TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT3847\_79240C.1.pep  
 3-10-424-599-262701

Query Match 95.0%; Score 19; DB 12; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 1.le+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
 ||||  
 5 VAEF 8

RESULT 60  
 3-10-363-616-368  
 Sequence 368, Application US/10363616  
 Publication No. US20040044181A1  
 GENERAL INFORMATION:  
 APPLICANT: Hyseq, Inc  
 TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
 FILE REFERENCE: 21272-113 (793)  
 CURRENT APPLICATION NUMBER: US/10/363,616  
 CURRENT FILING DATE: 2003-03-03  
 PRIOR APPLICATION NUMBER: 09/654,935  
 PRIOR FILING DATE: 2000-09-01  
 NUMBER OF SEQ ID NOS: 490  
 SEQ ID NO 368  
 LENGTH: 79  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 3-10-363-616-368

Query Match 95.0%; Score 19; DB 12; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 1.le+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
 ||||  
 24 VAEF 27

RESULT 61  
 3-10-424-599-216097  
 Sequence 216097, Application US/10424599  
 Publication No. US20040031072A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa Thomas J  
 APPLICANT: Kovalic David K  
 APPLICANT: Zhou Yihua  
 APPLICANT: Cao Yongwei  
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53223)B  
 CURRENT APPLICATION NUMBER: US/10/424,599  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 285684  
 SEQ ID NO 216097  
 LENGTH: 80  
 TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT3847\_37165C.1.pep  
 3-10-424-599-216097

Query Match 95.0%; Score 19; DB 12; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 1.le+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
 ||||  
 35 VAEF 38

RESULT 62  
 US-10-367-980A-14  
 Sequence 14, Application US/10367980A  
 Publication No. US2003022852A1  
 GENERAL INFORMATION:  
 APPLICANT: St Vincent's Institute of Medical Research  
 APPLICANT: Rogers, Suzanne D  
 APPLICANT: Best, James D  
 TITLE OF INVENTION: Human Facilitative Glucose Transport Protein GLUT8  
 FILE REFERENCE: VS:AJH:FP17928  
 CURRENT APPLICATION NUMBER: US/10/367,980A  
 CURRENT FILING DATE: 2003-02-19  
 PRIOR APPLICATION NUMBER: 09/509,731  
 PRIOR FILING DATE: 2003-02-19  
 NUMBER OF SEQ ID NOS: 42  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO 14  
 LENGTH: 82  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc.feature  
 OTHER INFORMATION: GLUT2  
 US-10-367-980A-14

Query Match 95.0%; Score 19; DB 15; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 1.le+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 ||||  
 DB 13 VAEF 16

RESULT 63  
 US-09-864-408A-5282  
 Sequence 5282, Application US/09864408A  
 Publication No. US20040009474A1  
 GENERAL INFORMATION:  
 APPLICANT: Leach, Martin D.  
 APPLICANT: Shinkets, Richard A.  
 TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod  
 FILE REFERENCE: 21402-012  
 CURRENT APPLICATION NUMBER: US/09/864,408A  
 CURRENT FILING DATE: 2001-05-24  
 PRIOR APPLICATION NUMBER: 60/206,690  
 PRIOR FILING DATE: 2000-05-24  
 NUMBER OF SEQ ID NOS: 9068  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 5282  
 LENGTH: 84  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-864-408A-5282

Query Match 95.0%; Score 19; DB 11; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 1.le+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 ||||  
 DB 52 VAEF 55

RESULT 64  
 US-10-078-090-125  
 Sequence 125, Application US/10078090  
 Publication No. US20030044815A1  
 GENERAL INFORMATION:  
 APPLICANT: Saiceda, Susana  
 APPLICANT: Macina, Roberto

```

/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 201853
/ LENGTH: 85
/ TYPE: PRT
/ ORGANISM: Glycine max
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(85)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_24299C.1.pep
US-10-424-599-201853

Query Match          95.0%; Score 19; DB 12; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
      ||||
Db      58 VAEF 61

RESULT 67
US-09-815-242-11224
/ Sequence 11224, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011a
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11224
/ LENGTH: 89
/ TYPE: PRT
/ ORGANISM: Haemophilus influenzae
US-09-815-242-11224

Query Match          95.0%; Score 19; DB 9; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
      ||||
Db      12 VAEF 15

```

## ESULT 68

S-09-815-242-11245  
 Sequence 11245, Application US/09815242  
 Patent No. US20020061569A1  
 GENERAL INFORMATION:  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari L.  
 APPLICANT: Zyskind, Judith W.  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John D.  
 APPLICANT: Carr, Grant J.  
 APPLICANT: Yamamoto, Robert T.  
 APPLICANT: Xu, H. Howard  
 TITLE OF INVENTION: Identification of Essential Genes in  
 FILE REFERENCE: ELITRA.011A  
 CURRENT APPLICATION NUMBER: US/09/815,242  
 CURRENT FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2001-02-16  
 NUMBER OF SEQ ID NOS: 14110  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 11245  
 LENGTH: 89  
 TYPE: PRT  
 ORGANISM: Haemophilus influenzae

Query Match 95.0%; Score 19; DB 9; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
 ||||  
 Db 12 VAEF 15

## ESULT 69

S-10-282-122A-58489  
 Sequence 58489, Application US/10282122A  
 Publication No. US20040029129A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari  
 APPLICANT: Zyskind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: ELITRA.034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 58489  
 ; LENGTH: 89  
 ; TYPE: PRT  
 ; ORGANISM: Haemophilus influenzae  
 US-10-282-122A-58489

Query Match 95.0%; Score 19; DB 12; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 ||||  
 Db 12 VAEF 15

## RESULT 70

US-10-282-122A-58530  
 Sequence 58530, Application US/10282122A  
 Publication No. US20040029129A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari  
 APPLICANT: Zyskind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: ELITRA.034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 58530  
; LENGTH: 89  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae

JS-10-282-122A-58530

Query Match 95.0%; Score 19; DB 12; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 VAEF 5  
Db 12 VAEF 15

## RESULT 71

US-10-282-122A-66960  
; Sequence 66960, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 66960  
; LENGTH: 89  
; TYPE: PRT  
; ORGANISM: Pasteurella multocida

US-10-282-122A-66960

Query Match 95.0%; Score 19; DB 12; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 2 VAEF 5  
Db 12 VAEF 15

## RESULT 72

US-10-282-122A-68774  
; Sequence 68774, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 68774  
; LENGTH: 89  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis

US-10-282-122A-68774

Query Match 95.0%; Score 19; DB 12; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 VAEF 5  
Db 12 VAEF 15

## RESULT 73

US-10-424-599-180307  
; Sequence 180307, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K

```
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 180307
LENGTH: 89
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_133832C.1.psp
-10-424-599-180307

Query Match          95.0%; Score 19; DB 12; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      2 VAEF 5
      ||||
      24 VAEF 27

;SULT 74
-10-097-111-360
Sequence 360, Application US/10097111
Publication No. US2003013877A1
GENERAL INFORMATION:
APPLICANT: PELLETIER, JERRY
APPLICANT: GROS, PHILLIPPE
APPLICANT: DUBOW, MICHAEL
TITLE OF INVENTION: DNA SEQUENCES FROM S. PNEUMONIAE BACTERIOPHAGE EPI THAT
FILE REFERENCE: 073406-0603
CURRENT APPLICATION NUMBER: US/10/097,111
CURRENT FILING DATE: 2002-07-24
PRIOR APPLICATION NUMBER: 09/676,412
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/157,218
PRIOR FILING DATE: 1999-09-30
NUMBER OF SEQ ID NOS: 552
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 360
LENGTH: 89
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
-10-097-111-360

Query Match          95.0%; Score 19; DB 14; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      2 VAEF 5
      ||||
      11 VAEF 14

;SULT 75
-10-424-599-245830
Sequence 245830, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
```

```
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 245830
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_64015C.1.psp
US-10-424-599-245830

Query Match          95.0%; Score 19; DB 12; Length 90;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      2 VAEF 5
      ||||
      60 VAEF 63

Search completed: May 24, 2004, 17:42:35
Job time : 41.6429 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

un on: May 24, 2004, 17:31:37 ; Search time 51.7857 Seconds  
(without alignments)

27.280 Million cell updates/sec

title: US-09-594-978A-1

perfect score: 20

sequence: 1 XVAEF 5

coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 1586107 seqs, 282547505 residues

total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

database : A\_Geneseq\_29Jan04.\*

1: Geneseqp1980s.\*

2: Geneseqp1980s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	19	95.0	6	5	AAM47151 S Chrysom
2	19	95.0	7	3	AAB07871 A beta-se
3	19	95.0	8	3	AAB07872 A beta-se
4	19	95.0	8	4	ABB56241 Vascular
5	19	95.0	8	4	AAB07870 DPI trypt
6	19	95.0	8	4	AAB07871 Schizophr
7	19	95.0	8	4	AAB07872 Schizophr
8	19	95.0	8	4	AAB07873 Schizophr
9	19	95.0	8	4	AAB07874 Schizophr
10	19	95.0	9	5	AAB07875 Schizophr
11	19	95.0	10	5	AAB07876 Schizophr
12	19	95.0	12	5	AAB06593 Beta-secr
13	19	95.0	12	5	AAB06592 Beta-secr
14	19	95.0	12	5	ABG78404 Memapsin
15	19	95.0	12	7	ADD35467 Escherich
16	19	95.0	13	3	AAB07889 A beta-se
17	19	95.0	13	4	AAB07888 A beta-se
18	19	95.0	14	3	AAB07887 A beta-se
19	19	95.0	14	4	AAB07886 A beta-se
20	19	95.0	15	6	AAB07885 A beta-se
21	19	95.0	20	2	AAB07884 A beta-se
22	19	95.0	20	2	AAB07883 A beta-se
23	19	95.0	21	4	AAB07882 A beta-se
24	19	95.0	21	4	AAB07881 A beta-se
25	19	95.0	25	2	AAB07880 A beta-se
26	19	95.0	25	4	AAB07879 A beta-se
27	19	95.0	25	4	AAB07878 A beta-se
28	19	95.0	25	4	AAB07877 A beta-se
29	19	95.0	25	4	AAB07876 A beta-se
30	19	95.0	25	4	AAB07875 A beta-se
31	19	95.0	25	4	AAB07874 A beta-se
32	19	95.0	25	4	AAB07873 A beta-se
33	19	95.0	25	4	AAB07872 A beta-se
34	19	95.0	25	4	AAB07871 A beta-se
35	19	95.0	25	4	AAB07870 A beta-se
36	19	95.0	25	4	AAB07869 A beta-se
37	19	95.0	25	4	AAB07868 A beta-se
38	19	95.0	25	4	AAB07867 A beta-se
39	19	95.0	25	4	AAB07866 A beta-se
40	19	95.0	25	4	AAB07865 A beta-se
41	19	95.0	25	4	AAB07864 A beta-se
42	19	95.0	25	4	AAB07863 A beta-se
43	19	95.0	25	4	AAB07862 A beta-se
44	19	95.0	25	4	AAB07861 A beta-se
45	19	95.0	25	4	AAB07860 A beta-se
46	19	95.0	25	4	AAB07859 A beta-se
47	19	95.0	25	4	AAB07858 A beta-se
48	19	95.0	25	4	AAB07857 A beta-se
49	19	95.0	25	4	AAB07856 A beta-se
50	19	95.0	25	4	AAB07855 A beta-se
51	19	95.0	25	4	AAB07854 A beta-se
52	19	95.0	25	4	AAB07853 A beta-se
53	19	95.0	25	4	AAB07852 A beta-se
54	19	95.0	25	4	AAB07851 A beta-se
55	19	95.0	25	4	AAB07850 A beta-se
56	19	95.0	25	4	AAB07849 A beta-se
57	19	95.0	25	4	AAB07848 A beta-se
58	19	95.0	25	4	AAB07847 A beta-se
59	19	95.0	25	4	AAB07846 A beta-se
60	19	95.0	25	4	AAB07845 A beta-se
61	19	95.0	25	4	AAB07844 A beta-se
62	19	95.0	25	4	AAB07843 A beta-se
63	19	95.0	25	4	AAB07842 A beta-se
64	19	95.0	25	4	AAB07841 A beta-se
65	19	95.0	25	4	AAB07840 A beta-se
66	19	95.0	25	4	AAB07839 A beta-se
67	19	95.0	25	4	AAB07838 A beta-se
68	19	95.0	25	4	AAB07837 A beta-se
69	19	95.0	25	4	AAB07836 A beta-se
70	19	95.0	25	4	AAB07835 A beta-se
71	19	95.0	25	4	AAB07834 A beta-se
72	19	95.0	25	4	AAB07833 A beta-se
73	19	95.0	25	4	AAB07832 A beta-se
74	19	95.0	25	4	AAB07831 A beta-se
75	19	95.0	25	4	AAB07830 A beta-se
76	19	95.0	25	4	AAB07829 A beta-se
77	19	95.0	25	4	AAB07828 A beta-se
78	19	95.0	25	4	AAB07827 A beta-se
79	19	95.0	25	4	AAB07826 A beta-se
80	19	95.0	25	4	AAB07825 A beta-se
81	19	95.0	25	4	AAB07824 A beta-se
82	19	95.0	25	4	AAB07823 A beta-se
83	19	95.0	25	4	AAB07822 A beta-se
84	19	95.0	25	4	AAB07821 A beta-se
85	19	95.0	25	4	AAB07820 A beta-se
86	19	95.0	25	4	AAB07819 A beta-se
87	19	95.0	25	4	AAB07818 A beta-se
88	19	95.0	25	4	AAB07817 A beta-se
89	19	95.0	25	4	AAB07816 A beta-se
90	19	95.0	25	4	AAB07815 A beta-se
91	19	95.0	25	4	AAB07814 A beta-se
92	19	95.0	25	4	AAB07813 A beta-se
93	19	95.0	25	4	AAB07812 A beta-se
94	19	95.0	25	4	AAB07811 A beta-se
95	19	95.0	25	4	AAB07810 A beta-se
96	19	95.0	25	4	AAB07809 A beta-se
97	19	95.0	25	4	AAB07808 A beta-se
98	19	95.0	25	4	AAB07807 A beta-se
99	19	95.0	25	4	AAB07806 A beta-se
100	19	95.0	25	4	AAB07805 A beta-se

26	19	95.0	28	2	AAW17832	AAW17832 Human mev
27	19	95.0	28	4	AAW33910	AAW33910 Peptide #
28	19	95.0	28	4	ABG55464	ABG55464 Human liv
29	19	95.0	28	4	ABG43601	ABG43601 Human pep
30	19	95.0	30	2	AAW62760	AAW62760 Streptoco
31	19	95.0	30	2	AAW82654	AAW82654 Human imm
32	19	95.0	33	4	AAO09024	AAO09024 Human pol
33	19	95.0	35	2	AAW77953	AAW77953 Antigenic
34	19	95.0	35	2	AAW46146	AAW46146 Predicted
35	19	95.0	35	2	AAW51751	AAW51751 H. influe
36	19	95.0	35	2	AAW55788	AAW55788 Human agu
37	19	95.0	35	2	AAW54505	AAW54505 Tbp2 anti
38	19	95.0	35	2	AAW54320	AAW54320 Human agu
39	19	95.0	35	3	AAW80448	AAW80448 H. influe
40	19	95.0	36	2	AAW27782	AAW27782 UDP-N-ace
41	19	95.0	37	4	AAW99833	AAW99833 Human exc
42	19	95.0	37	4	AAW42648	AAW42648 Human kid
43	19	95.0	43	4	AAW17544	AAW17544 Peptide #
44	19	95.0	43	4	ABB31362	ABB31362 Peptide #
45	19	95.0	43	4	ABB21904	ABB21904 Protein #
46	19	95.0	43	4	ABG51414	ABG51414 Human liv
47	19	95.0	43	4	AAW05214	AAW05214 Peptide #
48	19	95.0	47	3	AAW27178	AAW27178 RSV parti
49	19	95.0	48	4	AAW18541	AAW18541 Peptide #
50	19	95.0	48	4	ABB37578	ABB37578 Peptide #
51	19	95.0	48	4	AAW30996	AAW30996 Peptide #
52	19	95.0	48	4	ABB22868	ABB22868 Protein #
53	19	95.0	48	4	AAW70687	AAW70687 Human bon
54	19	95.0	48	4	ABG52389	ABG52389 Human liv
55	19	95.0	48	4	AAW06107	AAW06107 Peptide #
56	19	95.0	53	4	AAO10565	AAO10565 Human pol
57	19	95.0	53	4	AAO05247	AAO05247 Human pol
58	19	95.0	56	4	AAW87390	AAW87390 Human imm
59	19	95.0	56	6	ABP97126	ABP97126 Human mat
60	19	95.0	56	6	ABG76312	ABG76312 Human mat
61	19	95.0	58	2	AAW55888	AAW55888 Rat PCI2.
62	19	95.0	58	5	ABP04854	ABP04854 Human ORF
63	19	95.0	59	3	AAW60085	AAW60085 Arabidops
64	19	95.0	59	3	AAW60683	AAW60683 Arabidops
65	19	95.0	60	4	ABB42044	ABB42044 Peptide #
66	19	95.0	60	4	AAW35846	AAW35846 Peptide #
67	19	95.0	60	4	AAW75737	AAW75737 Human bon
68	19	95.0	60	4	AAW62925	AAW62925 Human bra
69	19	95.0	60	4	ABG57475	ABG57475 Human liv
70	19	95.0	62	2	AAW17262	AAW17262 H1H domai
71	19	95.0	62	4	AAW91247	AAW91247 Human imm
72	19	95.0	62	6	ABW65667	ABW65667 Propionib
73	19	95.0	64	4	AAO09517	AAO09517 Human pol
74	19	95.0	67	2	AAW48352	AAW48352 Human pro
75	19	95.0	68	5	ABP31047	ABP31047 Human ORF
76	19	95.0	69	3	AAW32809	AAW32809 Zea may
77	19	95.0	69	3	AAW21161	AAW21161 Zea may
78	19	95.0	69	4	ABB42591	ABB42591 Peptide #
79	19	95.0	69	4	AAW36821	AAW36821 Peptide #
80	19	95.0	69	4	AAW76713	AAW76713 Human bon
81	19	95.0	69	4	AAW63900	AAW63900 Human bra
82	19	95.0	69	4	ABG58405	ABG58405 Human liv
83	19	95.0	69	4	AAW91842	AAW91842 C glutami
84	19	95.0	71	5	ABW08064	ABW08064 Human ORF
85	19	95.0	71	5	ABW15877	ABW15877 Hepatitis
86	19	95.0	71	6	ADA35824	ADA35824 Acinetoba
87	19	95.0	73	4	AAU53554	AAU53554 Propionib
88	19	95.0	73	4	AAU66465	AAU66465 Propionib
89	19	95.0	73	5	ABP06455	ABP06455 Human ORF
90	19	95.0	73	6	ABW50073	ABW50073 Propionib
91	19	95.0	73	6	ABW62984	ABW62984 Propionib
92	19	95.0	74	5	AAE15876	AAE15876 Hepatitis
93	19	95.0	75	4	ABG00769	ABG00769 Novel hum
94	19	95.0	76	4	AAU45469	AAU45469 Propionib
95	19	95.0	76	6	ABM41988	ABM41988 Propionib
96	19	95.0	79	5	ABP62931	ABP62931 Human pol
97	19	95.0	83	6	ABP79418	ABP79418 N. gonorr
98	19	95.0	84	5	ABP66623	ABP66623 Human bre



99	19	95.0	84	5	ABP33668	Human ORF	172	19	95.0	133	4	ABB27760	Human pep
100	19	95.0	86	4	AU050485	Propionib	173	19	95.0	133	4	ABB18410	Protein #
101	19	95.0	86	6	ABM47004	Propionib	174	19	95.0	133	4	AA066115	Human bon
102	19	95.0	87	4	AA005209	Human pol	175	19	95.0	133	4	AA053732	Human bra
103	19	95.0	87	4	AAU43264	Propionib	176	19	95.0	133	4	ABG47783	Human liv
104	19	95.0	87	4	AU066125	Propionib	177	19	95.0	133	4	AA017071	Peptide #
105	19	95.0	87	6	ABM62644	Propionib	178	19	95.0	133	5	ABP10781	Human ORF
106	19	95.0	87	6	ABM39783	Propionib	179	19	95.0	133	5	ABG35765	Human pep
107	19	95.0	88	4	AU056968	Propionib	180	19	95.0	134	4	AU034920	Enterococ
108	19	95.0	88	6	ABM3487	Propionib	181	19	95.0	134	4	ABG02560	Novel hum
109	19	95.0	88	7	ADC89159	Ribosomal	182	19	95.0	135	5	ABP07711	Human ORF
110	19	95.0	89	3	ABM16761	Bacterioph	183	19	95.0	135	6	ABM69996	Photornab
111	19	95.0	89	4	AU035652	Haemophil	184	19	95.0	136	2	AAW55889	Rat plasmin
112	19	95.0	89	4	AU035631	Haemophil	185	19	95.0	139	3	AA010817	Methanoba
113	19	95.0	89	5	ABM54895	Lactococc	186	19	95.0	139	3	AA041153	Zea may
114	19	95.0	89	6	ABU40860	Protein e	187	19	95.0	139	4	ABM64375	Amino aci
115	19	95.0	89	6	ABU30365	Protein e	188	19	95.0	141	3	ABG02426	Human sec
116	19	95.0	89	6	ABU30606	Protein e	189	19	95.0	141	4	AAU02629	Propionib
117	19	95.0	89	6	ABU39036	Protein e	190	19	95.0	141	4	ABG30243	Novel hum
118	19	95.0	93	3	AA048060	Arabisdops	191	19	95.0	141	4	ABG07245	Novel hum
119	19	95.0	95	3	ABG34249	Arabisdops	192	19	95.0	141	6	ABM46788	Propionib
120	19	95.0	96	4	ABG07485	Novel hum	193	19	95.0	142	2	AAW10208	Canola co
121	19	95.0	97	5	ABM48132	Listeria	194	19	95.0	142	6	ABP79097	N. gonorr
122	19	95.0	97	6	ABP75366	Human sec	195	19	95.0	143	3	AA010816	Archaeogl
123	19	95.0	99	5	ABP41643	Human ova	196	19	95.0	143	5	AA014828	Human ste
124	19	95.0	100	3	AA0302653	Human sec	197	19	95.0	144	3	AA0505007	Arabisdops
125	19	95.0	100	4	AU062265	Propionib	198	19	95.0	144	2	AA034749	Chlamydia
126	19	95.0	100	6	ABM58784	Propionib	199	19	95.0	145	4	AAU27872	Human con
127	19	95.0	105	3	AA032808	Zea may	200	19	95.0	145	5	ABG98409	Haloarcul
128	19	95.0	106	3	AA032572	Arabisdops	201	19	95.0	145	5	ABJ11020	Yeast sel
129	19	95.0	106	5	ABM82190	P. multoc	202	19	95.0	145	5	ABJ11020	Yeast sel
130	19	95.0	107	3	AA061315	Arabisdops	203	19	95.0	146	6	ABP57361	Anti-TRAI
131	19	95.0	107	3	AA010318	Arabisdops	204	19	95.0	146	6	ABP57363	Anti-TRAI
132	19	95.0	107	4	AA030766	C glutami	205	19	95.0	146	6	ABU42336	Protein e
133	19	95.0	108	3	AA030023	Human sec	206	19	95.0	146	6	ABP56635	Ascaris s
134	19	95.0	109	4	ABG09614	Novel hum	207	19	95.0	146	7	AD26750	Human adi
135	19	95.0	109	5	ABM89162	Human pol	208	19	95.0	147	6	ABU07050	Maize SSI
136	19	95.0	109	5	ABP62930	Human pol	209	19	95.0	148	3	AA041152	Zea may
137	19	95.0	109	7	ABP99534	Mut-IL-4	210	19	95.0	148	6	ABP78900	N. gonorr
138	19	95.0	111	2	AA037827	Amino aci	211	19	95.0	148	7	ABM74182	DNA clone
139	19	95.0	111	3	AA0321160	Zea may	212	19	95.0	149	4	ABG05510	Novel hum
140	19	95.0	112	1	ABP90584	N-termina	213	19	95.0	149	5	ABM47889	Listeria
141	19	95.0	112	4	AA034813	Human pro	214	19	95.0	150	3	AA001131	Human sec
142	19	95.0	113	1	AA030583	N-termina	215	19	95.0	150	5	ABM49109	Listeria
143	19	95.0	113	3	AA030066	Human sec	216	19	95.0	152	5	AA052496	Superoxid
144	19	95.0	113	3	AA034248	Arabisdops	217	19	95.0	152	5	ABM51829	Human col
145	19	95.0	114	1	AA030953	N-termina	218	19	95.0	153	3	ABM40336	Human ORF
146	19	95.0	114	6	ABP97830	Amino aci	219	19	95.0	153	4	ABG07294	Novel hum
147	19	95.0	115	5	ABM06279	Human RS3	220	19	95.0	153	4	ABG10037	Novel hum
148	19	95.0	116	6	ABM64780	Propionib	221	19	95.0	153	4	ABG05309	Novel hum
149	19	95.0	117	5	ABP05042	Human ORF	222	19	95.0	154	3	AA041151	Zea may
150	19	95.0	119	3	AA033308	Pinus rad	223	19	95.0	154	3	AA038111	Arabisdops
151	19	95.0	119	6	ABM69462	Drosophil	224	19	95.0	154	3	AA042939	Arabisdops
152	19	95.0	122	4	ABP16105	Human ner	225	19	95.0	154	3	AA019886	Arabisdops
153	19	95.0	122	5	ABM31290	Human ORF	226	19	95.0	155	5	ABG80423	Arabisdops
154	19	95.0	125	4	AA031641	Human imm	227	19	95.0	156	4	ABM70411	Drosophil
155	19	95.0	125	4	AAU67165	Propionib	228	19	95.0	156	4	ABG24856	Novel hum
156	19	95.0	125	6	ABM63584	Propionib	229	19	95.0	156	4	ABM52504	Escherich
157	19	95.0	129	2	AA014573	Streptoco	230	19	95.0	157	4	ABM26258	Novel hum
158	19	95.0	129	3	AA030409	Arabisdops	231	19	95.0	157	4	ABG26258	Novel hum
159	19	95.0	131	4	AA030409	Arabisdops	232	19	95.0	159	2	AA011201	S. pneumo
160	19	95.0	131	4	AA030409	Arabisdops	233	19	95.0	162	3	AA062666	Arabisdops
161	19	95.0	131	4	AA030409	Arabisdops	234	19	95.0	162	6	ABM70110	Photornab
162	19	95.0	132	3	AA030335	Novel hum	235	19	95.0	163	3	AA012749	Arabisdops
163	19	95.0	132	4	AA030335	Novel hum	236	19	95.0	163	3	AA062665	Arabisdops
164	19	95.0	132	4	AA041985	Human pol	237	19	95.0	163	4	ABM63381	Drosophil
165	19	95.0	132	4	AA041985	Human pol	238	19	95.0	163	5	ABM97879	Human sec
166	19	95.0	132	4	AA041984	Human pol	239	19	95.0	164	2	AA029556	Human lun
167	19	95.0	133	3	AA041454	Human ORF	240	19	95.0	164	3	AA044479	Human lun
168	19	95.0	133	4	AA013986	Peptide #	241	19	95.0	164	4	AA013821	Human lun
169	19	95.0	133	4	AAU23379	Novel hum	242	19	95.0	164	5	ABM48330	Listeria
170	19	95.0	133	4	ABM2931	Peptide #	243	19	95.0	164	5	ABP31601	Human ORF
171	19	95.0	133	4	AA026392	Peptide #	244	19	95.0	164	7	AD066511	Human lun



391	19	95.0	224	6	ADA54376	Human pro	464	19	95.0	257	2	AAR97285	Aar97285 Human 265
392	19	95.0	224	6	ADA56787	Human sec	465	19	95.0	257	2	AAR21729	Aar21729 Nuclear m
393	19	95.0	224	7	ADC74046	Human sec	466	19	95.0	257	3	AAW59110	Aaw59110 Arabidops
394	19	95.0	224	7	ADD37843	Human sec	467	19	95.0	257	4	AAW74447	Aaw74447 Human col
395	19	95.0	226	4	AAW79581	Human pro	468	19	95.0	257	4	AAW92888	Aaw92888 C glutami
396	19	95.0	226	4	ABG03398	Novel hum	469	19	95.0	257	7	ADD27694	Add27694 Human adi
397	19	95.0	226	4	ABG03398	Novel hum	470	19	95.0	257	7	ADD26681	Add26681 Human adi
398	19	95.0	226	6	ADA90901	Novel hum	471	19	95.0	257	7	ADD27489	Add27489 Human adi
399	19	95.0	226	6	ADA36076	Acinetoba	472	19	95.0	258	3	AAW16704	Aaw16704 Bacteriop
400	19	95.0	227	5	AAE14342	Human pro	473	19	95.0	258	4	AAU36296	Aau36296 Pseudomon
401	19	95.0	228	3	AAV74793	Neisseria	474	19	95.0	258	4	AAU38502	Aau38502 Protein e
402	19	95.0	228	4	ABA7282	MST-37G2	475	19	95.0	258	6	ABU38502	Abu38502 Human pro
403	19	95.0	228	6	ABP78051	N. gonorr	476	19	95.0	260	4	AAU78433	Aau78433 Human pro
404	19	95.0	228	6	ABP80081	N. gonorr	477	19	95.0	260	4	AAU58456	Aau58456 Propionib
405	19	95.0	228	6	ABG75884	Mycobacte	478	19	95.0	260	4	ABG30334	Abg30334 Novel hum
406	19	95.0	229	4	AAW1949	S. epider	479	19	95.0	260	6	ABG54975	Abg54975 Propionib
407	19	95.0	229	4	AAU04888	Micromono	480	19	95.0	261	4	AAU00701	Aau00701 Mouse FCT
408	19	95.0	229	5	ABW1524	Herbicida	481	19	95.0	261	5	ABW79646	Abw79646 Mouse FCT
409	19	95.0	230	4	AAW80078	Corynebac	482	19	95.0	261	5	ABW78505	Abw78505 Murine FC
410	19	95.0	230	4	AAW90273	C glutami	483	19	95.0	261	5	ABW78505	Abw78505 Murine FC
411	19	95.0	231	2	AAW60988	Streptoco	484	19	95.0	261	6	ABU41453	Abu41453 Protein e
412	19	95.0	231	5	ABW50143	Listeria	485	19	95.0	261	6	ABU41453	Abu41453 Protein e
413	19	95.0	232	6	ADA36637	Acinetoba	486	19	95.0	262	4	ABG04663	Abg04663 Novel hum
414	19	95.0	232	3	AAV27368	Group B S	487	19	95.0	262	4	ABG04663	Abg04663 Novel hum
415	19	95.0	232	3	AAW11258	Arabidops	488	19	95.0	262	5	AAU72572	Aau72572 Arabidops
416	19	95.0	234	3	AAW22116	Maize glu	489	19	95.0	263	4	AAW95344	Aaw95344 Human adu
417	19	95.0	234	5	ABW1390	Herbicida	490	19	95.0	263	4	AAW90550	Aaw90550 Human sec
418	19	95.0	234	7	ADC96886	E. faeciu	491	19	95.0	263	5	ABG65464	Abg65464 Human alb
419	19	95.0	236	3	AAW42087	Human ORF	492	19	95.0	263	7	ADA44959	Ada44959 Human pol
420	19	95.0	236	4	AAW36145	Klebsiell	493	19	95.0	264	3	AAW15729	Aaw15729 Arabidops
421	19	95.0	236	4	AAW61537	Degp prot	494	19	95.0	264	4	ABG18250	Abg18250 Novel hum
422	19	95.0	236	4	ABG27447	Novel hum	495	19	95.0	265	2	AAW08269	Aaw08269 Mouse ser
423	19	95.0	238	3	ABW11695	Mouse ser	496	19	95.0	265	2	AAW70534	Aaw70534 Mouse neu
424	19	95.0	238	3	AAW28132	Glycerald	497	19	95.0	265	6	ABU24490	Abu24490 Protein e
425	19	95.0	238	4	AAW61538	Degp prot	498	19	95.0	266	4	AAW79417	Aaw79417 Human pro
426	19	95.0	238	4	AAW83029	S. epider	499	19	95.0	267	2	AAW85606	Aaw85606 Secreted
427	19	95.0	238	4	AAW83016	S. epider	500	19	95.0	267	2	ABU50254	Abu50254 Protein e
428	19	95.0	238	5	ABW53557	Lactococc	501	19	95.0	268	5	AAE15256	Aae15256 Human RNA
429	19	95.0	238	5	AAU80750	Eucalyptu	502	19	95.0	268	5	ABP61845	Abp61845 Human pol
430	19	95.0	238	6	ABU44475	Protein e	503	19	95.0	269	2	AAW55786	Aaw55786 Human aqu
431	19	95.0	238	6	ABU42738	Protein e	504	19	95.0	269	2	AAW94322	Aaw94322 Human aqu
432	19	95.0	238	6	ABU42738	Protein e	505	19	95.0	269	4	ABG26217	Abg26217 Novel hum
433	19	95.0	238	6	ABM72778	Staphyloc	506	19	95.0	269	6	ABP53586	Abp53586 Protein s
434	19	95.0	239	3	AAV74795	Neisseria	507	19	95.0	269	6	ABP58050	Abp58050 Human equ
435	19	95.0	239	3	AAW11257	Arabidops	508	19	95.0	269	6	ABP64028	Abp64028 Human pro
436	19	95.0	239	3	AAW47494	Neisseria	509	19	95.0	270	4	AAW90189	Aaw90189 C glutami
437	19	95.0	239	3	AAW96778	Putative	510	19	95.0	270	6	ADA05788	Ada05788 Human NOV
438	19	95.0	240	3	AAW19884	Arabidops	511	19	95.0	270	6	ADC08133	Adc08133 Rice prot
439	19	95.0	240	3	AAW42937	Arabidops	512	19	95.0	271	2	AAV15227	Aav15227 Human rec
440	19	95.0	240	7	AAW25064	Plant gro	513	19	95.0	271	3	AAW28205	Aaw28205 Novel hum
441	19	95.0	241	5	ABP40753	Staphyloc	514	19	95.0	271	3	AAW31287	Aaw31287 Arabidops
442	19	95.0	241	6	ABU28310	Protein e	515	19	95.0	271	3	AAW22050	Aaw22050 Arabidops
443	19	95.0	243	2	AAW96089	S. pneumo	516	19	95.0	271	4	AAW92851	Aaw92851 Human pro
444	19	95.0	243	5	ABP38742	Staphyloc	517	19	95.0	271	4	AAW88446	Aaw88446 Human mem
445	19	95.0	244	5	ABW33337	Human TSA	518	19	95.0	271	4	AAW64357	Aaw64357 Human sig
446	19	95.0	244	6	ABU4263	Protein e	519	19	95.0	271	5	ABG90282	Abg90282 Human pol
447	19	95.0	245	4	ABU4263	Protein e	520	19	95.0	271	5	AAW90282	Aaw90282 Human pol
448	19	95.0	245	7	ADC08337	Novel hum	521	19	95.0	271	7	ADC95444	Adc95444 E. faeciu
449	19	95.0	247	7	AAW11660	Rice prot	522	19	95.0	272	4	AAW23799	Aaw23799 Novel hum
450	19	95.0	248	2	AAW11660	S. pneumo	523	19	95.0	272	4	AAW94302	Aaw94302 Human rep
451	19	95.0	249	4	AAW19896	Human TAN	524	19	95.0	272	6	ABU42431	Abu42431 Protein e
452	19	95.0	249	6	ABU11227	Polyketid	525	19	95.0	272	6	ABM71001	Abm71001 Staphyloc
453	19	95.0	249	6	ABU11227	Polyketid	526	19	95.0	272	6	ABM71001	Abm71001 Staphyloc
454	19	95.0	250	2	AAW61593	Z. mays a	527	19	95.0	273	3	AAW15728	Aaw15728 Arabidops
455	19	95.0	250	2	AAW33913	Sequence	528	19	95.0	273	3	AAW15728	Aaw15728 Arabidops
456	19	95.0	250	2	AAW60153	Nematode-	529	19	95.0	273	6	ABR42553	AbR42553 Simocycli
457	19	95.0	250	2	AAW7082	Tobacco r	530	19	95.0	273	6	ABR42553	AbR42553 Simocycli
458	19	95.0	250	3	AAW42541	Human ORF	531	19	95.0	273	7	AAW08633	Aaw08633 Novel pro
459	19	95.0	250	3	AAW42541	Human ORF	532	19	95.0	273	7	AAW08633	Aaw08633 Novel pro
460	19	95.0	250	6	AAW47372	Propionib	533	19	95.0	274	3	AAW00211	Aaw00211 Putative
461	19	95.0	250	6	AAW47372	Propionib	534	19	95.0	274	3	AAW00211	Aaw00211 Putative
462	19	95.0	253	4	AAW34686	E. coli c	535	19	95.0	274	4	AAW63322	Aaw63322 Human bre
463	19	95.0	253	7	ADC95445	E. faeciu	536	19	95.0	274	6	ABU23635	Abu23635 Protein e
	19	95.0	254	6	ADA05790	Human NOV		19	95.0	274	6	ABU46447	Abu46447 Protein e
	19	95.0	256	6	ABM67840	Phototrab		19	95.0	274	7	ADC32869	Adc32869 Human nov

537	19	95.0	275	6	ABU24121	Abu24121 Protein e	610	19	95.0	300	7	ADC31643	Adc31643 Human nov
538	19	95.0	276	3	AAU42611	Aau42611 Human ORF	611	19	95.0	301	4	AAG82151	Aag82151 S. epider
539	19	95.0	277	3	AAW18875	Aaw18875 Arabesia e	612	19	95.0	302	3	AAG15595	Aag15595 Arabidops
540	19	95.0	277	3	AAG22852	Aag22852 Arabidops	613	19	95.0	303	3	AAG20322	Aag20322 Arabidops
541	19	95.0	277	3	AAG04958	Aag04958 Arabidops	614	19	95.0	303	4	AAG89980	C glutami
542	19	95.0	278	3	AAG07719	Aag07719 Arabidops	615	19	95.0	304	3	AAG37633	Arabidops
543	19	95.0	278	3	AAG39551	Aag39551 Arabidops	616	19	95.0	304	3	AAG04957	Arabidops
544	19	95.0	278	3	AAG37634	Aag37634 Arabidops	617	19	95.0	304	3	AAG30063	Arabidops
545	19	95.0	278	5	ABP30329	Abp30329 Streptoco	618	19	95.0	304	3	AAG30713	Arabidops
546	19	95.0	278	6	ABP77659	Abp77659 N. gonorr	619	19	95.0	304	3	AAG39550	Arabidops
547	19	95.0	278	6	ABU37537	Abu37537 Protein e	620	19	95.0	305	3	AAG30713	Arabidops
548	19	95.0	278	6	ABU37837	Abu37837 Protein e	621	19	95.0	305	3	AAG10620	Arabidops
549	19	95.0	279	3	AAG54754	Aag54754 Arabidops	622	19	95.0	305	4	AAG92837	C glutami
550	19	95.0	279	4	ABG30336	Abg30336 Novel hum	623	19	95.0	306	3	AAG40089	Arabidops
551	19	95.0	279	5	ABBS5091	Abbs5091 Lactococc	624	19	95.0	306	3	AAG13663	Arabidops
552	19	95.0	280	2	AAW03521	Aaw03521 Bali modi	625	19	95.0	306	4	AU03620	Group B S
553	19	95.0	280	2	AAV37073	Aav37073 Amino aci	626	19	95.0	306	5	ABP29888	Streptoco
554	19	95.0	280	4	AAG30751	Aag30751 C glutami	627	19	95.0	306	5	ABP28497	Streptoco
555	19	95.0	281	3	AAG26675	Aag26675 Arabidops	628	19	95.0	306	7	ADB70040	C. neofo
556	19	95.0	281	5	ABG77963	Abg77963 Ehrlichia	629	19	95.0	307	7	ADC37261	Nuclear f
557	19	95.0	281	6	ADA09791	Ada09791 E. canis	630	19	95.0	308	3	AB53390	Human col
558	19	95.0	282	2	AAW34205	Aaw34205 Streptomy	631	19	95.0	308	4	AAG92374	C glutami
559	19	95.0	282	2	AAW55806	Aaw55806 Streptomy	632	19	95.0	308	5	ABBS4256	Lactococc
560	19	95.0	282	3	AAW57116	Aaw57116 Human pro	633	19	95.0	308	5	ABP60904	Lactococc
561	19	95.0	282	3	AAG13664	Aag13664 Arabidops	634	19	95.0	308	5	ABP65647	Bifidobac
562	19	95.0	282	3	AAG40090	Aag40090 Arabidops	635	19	95.0	308	7	ADC79410	Human G p
563	19	95.0	282	6	ABM67054	Abm67054 Photorhab	636	19	95.0	309	3	ADC79410	Human G p
564	19	95.0	282	6	ABU19351	Abu19351 Protein e	637	19	95.0	309	5	ABR93102	Herbicida
565	19	95.0	283	2	AAW89755	Aaw89755 Staphyloc	638	19	95.0	309	5	ABP61176	Human nuc
566	19	95.0	284	4	AAW39850	Aaw39850 Human pro	639	19	95.0	310	3	AAE21021	Human nuc
567	19	95.0	285	2	AAW95957	Aaw95957 Eukaryoti	640	19	95.0	310	5	AAW51006	Corynebac
568	19	95.0	285	6	ABU21606	Abu21606 Protein e	641	19	95.0	310	6	AAG79924	Human Sir
569	19	95.0	285	6	ABU41998	Abu41998 Protein e	642	19	95.0	310	6	ABU39481	Protein e
570	19	95.0	286	6	ADA05784	Ada05784 Human NOV	643	19	95.0	311	3	AB11697	Mouse ser
571	19	95.0	286	6	ABU41392	Abu41392 Protein e	644	19	95.0	311	3	AAG46265	Arabidops
572	19	95.0	287	3	AAG26674	Aag26674 Arabidops	645	19	95.0	311	4	AAE11899	Human G-p
573	19	95.0	287	3	ABU00416	Abu00416 Human nov	646	19	95.0	311	4	AAE71742	Human olf
574	19	95.0	288	6	ABM69140	Abm69140 Photorhab	647	19	95.0	311	4	AAG72431	Human olf
575	19	95.0	288	6	ABU38998	Abu38998 Protein e	648	19	95.0	311	4	AAG71759	Human OR-
576	19	95.0	289	4	AAU35600	Aau35600 Haemophil	649	19	95.0	311	4	AAE10690	G-protein
577	19	95.0	289	6	ABU30505	Abu30505 Protein e	650	19	95.0	311	4	AAG89917	C glutami
578	19	95.0	290	3	AAW78511	Aaw78511 Human unc	651	19	95.0	311	4	AAU24520	Human olf
579	19	95.0	290	3	AAG23887	Aag23887 Arabidops	652	19	95.0	311	5	ABP95725	Human GPC
580	19	95.0	290	6	ADA05786	Ada05786 Human NOV	653	19	95.0	311	5	ABP95725	Human GPC
581	19	95.0	291	5	ABB90394	Abb90394 Human pol	654	19	95.0	311	5	ABP95777	Human GPC
582	19	95.0	291	7	ADC31166	Adc31166 Human nov	655	19	95.0	311	5	AAU95637	Human olf
583	19	95.0	291	7	ADC87617	Adc87617 Human GPC	656	19	95.0	311	5	AAU95535	Human olf
584	19	95.0	292	3	AAG26171	Aag26171 Arabidops	657	19	95.0	311	5	ABJ04727	GPCR 7 pr
585	19	95.0	292	3	AAG60250	Aag60250 Arabidops	658	19	95.0	311	5	ABJ04727	GPCR 7 pr
586	19	95.0	293	6	ABU48542	Abu48542 Protein e	659	19	95.0	311	5	ABJ04727	GPCR 7 pr
587	19	95.0	294	5	ABG91459	Abg91459 Purine/py	660	19	95.0	311	5	ABJ04727	GPCR 7 pr
588	19	95.0	294	6	ABR53328	Abx53328 Protein s	661	19	95.0	311	6	ABU11179	Human G-p
589	19	95.0	294	6	ABP98029	Abp98029 Protein p	662	19	95.0	311	6	ABU11195	Human G-p
590	19	95.0	295	3	AAG31286	Aag31286 Arabidops	663	19	95.0	311	7	ADC85801	Human GPC
591	19	95.0	295	5	ABP28274	Abp28274 Streptoco	664	19	95.0	311	7	ADC85801	Human GPC
592	19	95.0	295	5	ABG91460	Abg91460 Purine/py	665	19	95.0	311	7	ADD12777	Novel hum
593	19	95.0	295	5	ABG91461	Abg91461 Purine/py	666	19	95.0	312	3	AAG39549	Arabidops
594	19	95.0	295	5	ABG91449	Abg91449 Aromatic	667	19	95.0	312	3	AAG04956	Arabidops
595	19	95.0	296	5	AAW52345	Aaw52345 Protein e	668	19	95.0	312	3	AAG04956	Arabidops
596	19	95.0	297	6	ABU22319	Abu22319 Protein e	669	19	95.0	312	3	AAG07717	Arabidops
597	19	95.0	297	6	ABU22270	Abu22270 Protein e	670	19	95.0	312	3	AAG07717	Arabidops
598	19	95.0	299	2	AAW06968	Aaw06968 E. canis	671	19	95.0	312	4	AAG72585	Human OR-
599	19	95.0	299	3	AAW07975	Aaw07975 A corn tr	672	19	95.0	312	4	AAG72585	Human OR-
600	19	95.0	299	4	AAW33403	Aaw33403 Human pro	673	19	95.0	312	4	AAE06767	Human G-p
601	19	95.0	299	4	AAW87731	Aaw87731 Human t2R	674	19	95.0	312	4	AAE06767	Human G-p
602	19	95.0	299	5	ABG77957	Abg77957 Ehrlichia	675	19	95.0	312	4	AAE06767	Human G-p
603	19	95.0	299	6	ABR38899	Abx38899 Human hr2	676	19	95.0	312	5	ABR06624	G protein
604	19	95.0	299	6	ADA09779	Ada09779 E. canis	677	19	95.0	312	5	ABR06624	G protein
605	19	95.0	299	7	ADC97734	Adc97734 Rhizobio	678	19	95.0	312	5	ABP95923	Human GPC
606	19	95.0	300	3	AAG45274	Aag45274 Arabidops	679	19	95.0	312	5	ABP95923	Human GPC
607	19	95.0	300	3	AAG30714	Aag30714 Arabidops	680	19	95.0	312	5	AAU95531	Human olf
608	19	95.0	300	3	AAG23886	Aag23886 Arabidops	681	19	95.0	312	5	AAU95531	Human olf
609	19	95.0	300	5	ABP27828	Abp27828 Streptoco	682	19	95.0	312	6	ABU26073	Protein s

683	19	95.0	312	7	ADC85589	Adc85589 Human GPC	756	19	95.0	328	4	AAU35342	Aau35342 Enterococ
684	19	95.0	312	7	ADD13513	Add13513 C. glutam	757	19	95.0	328	5	ABB91780	Abb91780 Herbicida
685	19	95.0	313	2	AAR37346	Aar37346 PEP PM. 1	758	19	95.0	328	6	ABO23285	Abo23285 Mouse ary
686	19	95.0	313	5	ABR49170	Abr49170 Listeria	759	19	95.0	328	6	ABO23284	Abo23284 Cow aryl-
687	19	95.0	314	3	AAV99818	Aav99818 Wheat par	760	19	95.0	328	6	ABU28954	Abu28954 Protein e
688	19	95.0	314	6	ABR01619	Abr01619 Human G p	761	19	95.0	328	6	ABM73204	Abm73204 Staphyloc
689	19	95.0	314	6	ABR011585	Abr011585 Human G p	762	19	95.0	329	5	AAW80623	Aaw80623 S. pneumo
690	19	95.0	315	4	ABU17461	Abu17461 Protein e	763	19	95.0	329	5	ABP27977	Abp27977 Streptoco
691	19	95.0	315	4	ABBS2707	Abbs2707 Escherich	764	19	95.0	329	7	ABM73892	Abm73892 DNA clone
692	19	95.0	315	4	AGG71585	Agg71585 Human olf	765	19	95.0	331	2	AAV37504	Aav37504 Protein i
693	19	95.0	315	4	AGG72285	Agg72285 Human olf	766	19	95.0	331	2	AAV37504	Aav37504 Protein i
694	19	95.0	315	4	AGG72890	Agg72890 Human olf	767	19	95.0	331	4	AAU36168	Aau36168 Klebsiell
695	19	95.0	315	5	AAE17492	Aae17492 Human sec	768	19	95.0	331	4	AAV72652	Aav72652 Murine OR
696	19	95.0	315	5	AAE17213	Aae17213 Human GPC	769	19	95.0	334	2	AAV36939	Aav36939 Chlamydia
697	19	95.0	315	6	ABM70746	Abm70746 Photorhab	770	19	95.0	334	4	AAE06056	Aae06056 Human gen
698	19	95.0	315	6	ABR01616	Abr01616 Human G p	771	19	95.0	334	5	ABF41314	Abf41314 Human ova
699	19	95.0	315	7	ADC37263	Adc37263 Nuclear f	772	19	95.0	334	5	ABG33878	Abg33878 Human sec
700	19	95.0	315	7	ADD12827	Add12827 Novel hum	773	19	95.0	334	6	ADA57046	Ada57046 Human sec
701	19	95.0	315	7	ADG90061	Adg90061 Human rho	774	19	95.0	334	6	ADA40900	Ada40900 Human sec
702	19	95.0	316	3	AAG20321	Aag20321 Arabidops	775	19	95.0	334	6	ABU43221	Abu43221 Protein e
703	19	95.0	316	3	AAU55622	Aau55622 Propionib	776	19	95.0	334	6	ABR61839	Abr61839 Arabidops
704	19	95.0	316	4	ABP98031	Abp98031 Protein p	777	19	95.0	335	2	AAV57736	Aav57736 Hyperseus
705	19	95.0	316	6	ABW52141	Abw52141 Propionib	778	19	95.0	335	3	AAV87079	Aav87079 Human sec
706	19	95.0	317	2	AAV011381	Aav011381 Rat osteo	779	19	95.0	336	3	AAV46263	Aav46263 Arabidops
707	19	95.0	317	2	AAE41562	Aae41562 Human ORF	780	19	95.0	336	3	AAV32249	Aav32249 Arabidops
708	19	95.0	317	3	ABM69850	Abm69850 Photorhab	781	19	95.0	336	3	ABG11256	Abg11256 Arabidops
709	19	95.0	317	7	ADE57454	Ade57454 Rat Prote	782	19	95.0	336	6	ABU38643	Abu38643 Protein e
710	19	95.0	318	2	AAW21746	Aaw21746 E3330-bin	783	19	95.0	336	6	ABU50329	Abu50329 Protein e
711	19	95.0	318	2	AAW48894	Aaw48894 Homo sapi	784	19	95.0	337	3	AAV67495	Aav67495 Arabidops
712	19	95.0	318	2	AAW52864	Aaw52864 Apurinic/	785	19	95.0	337	3	AAV43811	Aav43811 Arabidops
713	19	95.0	318	3	AAV30062	Aav30062 Arabidops	786	19	95.0	337	3	AAV06421	Aav06421 Arabidops
714	19	95.0	318	3	AAV08737	Aav08737 Arabidops	787	19	95.0	337	3	AAV40549	Aav40549 Arabidops
715	19	95.0	318	3	AAV34921	Aav34921 Arabidops	788	19	95.0	338	2	ADA35864	Ada35864 Acinetoba
716	19	95.0	318	5	ABP60935	Abp60935 Zymomonas	789	19	95.0	339	2	AAV07063	Aav07063 Renal can
717	19	95.0	318	7	ADD08904	Add08904 Apurinic/	790	19	95.0	339	4	ABM61273	Abm61273 Human TAN
718	19	95.0	318	7	ADG25727	Adg25727 Human pro	791	19	95.0	339	4	ABM61273	Abm61273 Human TAN
719	19	95.0	319	3	AAV31049	Aav31049 Arabidops	792	19	95.0	339	4	ABM61274	Abm61274 Human TAN
720	19	95.0	319	4	ABG61257	Abg61257 Mature hu	793	19	95.0	339	4	ABM61275	Abm61275 Human TAN
721	19	95.0	319	4	ABM49403	Abm49403 Human gly	794	19	95.0	339	4	AAV61276	Aav61276 Human TAN
722	19	95.0	319	6	ABP98028	Abp98028 Protein p	795	19	95.0	339	4	AAV72790	Aav72790 Human pla
723	19	95.0	319	6	ABU11223	Abu11223 Human TAN	796	19	95.0	339	4	AAV38760	Aav38760 Human pol
724	19	95.0	320	3	AAU50958	Aau50958 Arabidops	797	19	95.0	339	4	ABM31668	Abm31668 Amino aci
725	19	95.0	320	3	AAV06286	Aav06286 Arabidops	798	19	95.0	339	4	ABM31668	Abm31668 Amino aci
726	19	95.0	320	3	AAV40088	Aav40088 Arabidops	799	19	95.0	339	6	ABP72608	Abp72608 Human pro
727	19	95.0	320	3	AAV13662	Aav13662 Arabidops	800	19	95.0	339	6	ABP72607	Abp72607 Benign pr
728	19	95.0	320	4	AAV2947	Aav2947 S. epider	801	19	95.0	339	6	ADA84105	Ada84105 Human GP6
729	19	95.0	320	5	ABM91676	Abm91676 Herbicida	802	19	95.0	339	6	ABU11221	Abu11221 Human TAN
730	19	95.0	321	3	AAV99817	Aav99817 Soybean p	803	19	95.0	339	6	ABU11221	Abu11221 Human TAN
731	19	95.0	321	5	ABP51575	Abp51575 Human G p	804	19	95.0	339	6	ABU11240	Abu11240 Glycoprot
732	19	95.0	321	6	ADA89479	Ada89479 Staphyloc	805	19	95.0	339	6	ABU11242	Abu11242 Glycoprot
733	19	95.0	322	3	AAV94666	Aav94666 Human unc	806	19	95.0	339	6	ABU11242	Abu11242 Glycoprot
734	19	95.0	322	3	AAV94668	Aav94668 Murine un	807	19	95.0	339	6	ABU70998	Abu70998 Human adi
735	19	95.0	322	3	AAV20529	Aav20529 Arabidops	808	19	95.0	339	6	ABR82451	Abr82451 Human ARP
736	19	95.0	322	3	AAV08879	Aav08879 Arabidops	809	19	95.0	340	6	ABR42679	Abr42679 Pseudomon
737	19	95.0	322	6	ABU25221	Abu25221 Protein e	810	19	95.0	341	4	ABG05957	Abg05957 Novel hum
738	19	95.0	322	6	ABM72033	Abm72033 Staphyloc	811	19	95.0	341	4	ABG05957	Abg05957 Novel hum
739	19	95.0	323	3	AAV06285	Aav06285 Arabidops	812	19	95.0	342	4	ABM11919	Abm11919 Human cyt
740	19	95.0	323	6	ABU41982	Abu41982 Protein e	813	19	95.0	342	4	ABM11919	Abm11919 Human cyt
741	19	95.0	324	3	AAV06284	Aav06284 Arabidops	814	19	95.0	342	5	ABG15727	Abg15727 Arabidops
742	19	95.0	324	6	ABU33384	Abu33384 Protein e	815	19	95.0	343	3	ABG45273	Abg45273 Arabidops
743	19	95.0	324	6	ABU29228	Abu29228 Protein e	816	19	95.0	343	5	ABM1220	Abm1220 Herbicida
744	19	95.0	324	6	ADA333349	Ada333349 Acinetoba	817	19	95.0	343	5	ABM1220	Abm1220 Herbicida
745	19	95.0	325	2	AAV50034	Aav50034 Porcine c	818	19	95.0	344	3	ABM22851	Abm22851 Arabidops
746	19	95.0	325	3	AAV94669	Aav94669 Murine un	819	19	95.0	345	3	ABM56610	Abm56610 Human pro
747	19	95.0	325	3	AAV94665	Aav94665 Human unc	820	19	95.0	345	7	ADC94222	Adc94222 E. faeciu
748	19	95.0	325	3	AAV03320	Aav03320 Arabidops	821	19	95.0	346	3	AAV32505	Aav32505 S. lavend
749	19	95.0	325	6	ADA05732	Ada05732 Human NOV	822	19	95.0	346	4	AAV40546	Aav40546 Human pol
750	19	95.0	326	3	AAV23993	Aav23993 Hepatitis	823	19	95.0	347	7	ADE10277	Adel10277 S. lavend
751	19	95.0	326	3	AAV46264	Aav46264 Arabidops	824	19	95.0	347	4	ABG93168	Abg93168 C glutami
752	19	95.0	326	3	AAV51053	Aav51053 Arabidops	825	19	95.0	347	5	ABP38453	Abp38453 Staphyloc
753	19	95.0	326	5	ABG31930	Abg31930 Herbicida	826	19	95.0	347	5	ABP41750	Abp41750 Human ova
754	19	95.0	327	2	AAV50033	Aav50033 Porcine c	827	19	95.0	347	7	ADC54077	Adc54077 Mouse tes
755	19	95.0	327	6	ABU11347	Abu11347 Protein e	828	19	95.0	347	7	ADC54079	Adc54079 Mouse tes

829	19	95.0	350	3	ABO2086	Aab02086 F420-depe	902	19	95.0	377	6	ABU25867	Abu25867 Protein e
830	19	95.0	350	5	AAO15202	Aao15202 Rhodococc	903	19	95.0	378	4	AAW53361	Aam93361 Human pol
831	19	95.0	350	5	AAU97147	Aau97147 Rhodococc	904	19	95.0	378	5	AAE20087	Aae20087 Laccobaci
832	19	95.0	350	5	AAE16748	Aae16748 Rhodococc	905	19	95.0	379	6	ABU23947	Abu23947 Protein e
833	19	95.0	351	6	ABU18964	Abu18964 Protein e	906	19	95.0	379	6	ABU49216	Abu49216 Protein e
834	19	95.0	352	5	ABBA48732	Abba48732 Listeria	907	19	95.0	379	6	ABM04817	Abm04817 Rat cdc 3
835	19	95.0	352	6	ABU32749	Abu32749 Protein e	908	19	95.0	379	7	ADD45256	Add45256 Rat Prote
836	19	95.0	353	3	AAU94667	Aau94667 Human unc	909	19	95.0	380	2	AAW98800	Aaw98800 H. pylori
837	19	95.0	353	3	AAU94667	Aau94667 Human unc	910	19	95.0	380	2	AAW98800	Aaw98800 H. pylori
838	19	95.0	353	3	AAU94667	Aau94667 Human unc	911	19	95.0	380	4	AAU35898	Aau35898 Helicobac
839	19	95.0	353	3	AAU94667	Aau94667 Human unc	912	19	95.0	380	4	AAU35898	Aau35898 Helicobac
840	19	95.0	354	3	AAU94667	Aau94667 Human unc	913	19	95.0	380	4	AAU35898	Aau35898 Helicobac
841	19	95.0	354	3	AAU94667	Aau94667 Human unc	914	19	95.0	380	4	AAU35898	Aau35898 Helicobac
842	19	95.0	356	6	ADA36384	Ada36384 Acinetoba	915	19	95.0	383	6	AAU5985	Aau5985 Protein e
843	19	95.0	357	3	AAU94667	Aau94667 Human unc	916	19	95.0	383	6	AAU5985	Aau5985 Protein e
844	19	95.0	357	3	AAU94667	Aau94667 Human unc	917	19	95.0	383	6	AAU5985	Aau5985 Protein e
845	19	95.0	357	3	AAU94667	Aau94667 Human unc	918	19	95.0	384	3	AAU71056	Aau71056 Human mem
846	19	95.0	357	3	AAU94667	Aau94667 Human unc	919	19	95.0	384	3	AAU71056	Aau71056 Human mem
847	19	95.0	357	3	AAU94667	Aau94667 Human unc	920	19	95.0	384	3	AAU71056	Aau71056 Human mem
848	19	95.0	358	3	AAU94667	Aau94667 Human unc	921	19	95.0	384	4	AAU40199	Aam40199 Human pol
849	19	95.0	358	3	AAU94667	Aau94667 Human unc	922	19	95.0	384	4	AAU40199	Aam40199 Human pol
850	19	95.0	359	3	AAU94667	Aau94667 Human unc	923	19	95.0	384	4	AAU40199	Aam40199 Human pol
851	19	95.0	359	3	AAU94667	Aau94667 Human unc	924	19	95.0	384	4	AAU40199	Aam40199 Human pol
852	19	95.0	360	4	AAU40979	Aam40979 Human col	925	19	95.0	384	4	AAU40199	Aam40199 Human pol
853	19	95.0	360	4	AAU40979	Aam40979 Human col	926	19	95.0	384	4	AAU40199	Aam40199 Human pol
854	19	95.0	361	7	ADE09991	Ade09991 Novel pro	927	19	95.0	385	5	AAU72542	Aau72542 Arabidops
855	19	95.0	361	7	ADE09991	Ade09991 Novel pro	928	19	95.0	385	5	AAU72542	Aau72542 Arabidops
856	19	95.0	361	7	ADE09991	Ade09991 Novel pro	929	19	95.0	386	3	AAU72542	Aau72542 Arabidops
857	19	95.0	361	7	ADE09991	Ade09991 Novel pro	930	19	95.0	386	3	AAU72542	Aau72542 Arabidops
858	19	95.0	361	7	ADE09991	Ade09991 Novel pro	931	19	95.0	386	3	AAU72542	Aau72542 Arabidops
859	19	95.0	369	4	ABG01143	Abg01143 Novel hum	932	19	95.0	386	3	AAU72542	Aau72542 Arabidops
860	19	95.0	369	4	ABG01143	Abg01143 Novel hum	933	19	95.0	386	3	AAU72542	Aau72542 Arabidops
861	19	95.0	369	4	ABG01143	Abg01143 Novel hum	934	19	95.0	386	3	AAU72542	Aau72542 Arabidops
862	19	95.0	369	4	ABG01143	Abg01143 Novel hum	935	19	95.0	386	3	AAU72542	Aau72542 Arabidops
863	19	95.0	369	4	ABG01143	Abg01143 Novel hum	936	19	95.0	386	3	AAU72542	Aau72542 Arabidops
864	19	95.0	369	4	ABG01143	Abg01143 Novel hum	937	19	95.0	386	3	AAU72542	Aau72542 Arabidops
865	19	95.0	369	4	ABG01143	Abg01143 Novel hum	938	19	95.0	386	3	AAU72542	Aau72542 Arabidops
866	19	95.0	369	4	ABG01143	Abg01143 Novel hum	939	19	95.0	386	3	AAU72542	Aau72542 Arabidops
867	19	95.0	369	4	ABG01143	Abg01143 Novel hum	940	19	95.0	386	3	AAU72542	Aau72542 Arabidops
868	19	95.0	369	4	ABG01143	Abg01143 Novel hum	941	19	95.0	386	3	AAU72542	Aau72542 Arabidops
869	19	95.0	369	4	ABG01143	Abg01143 Novel hum	942	19	95.0	386	3	AAU72542	Aau72542 Arabidops
870	19	95.0	369	4	ABG01143	Abg01143 Novel hum	943	19	95.0	386	3	AAU72542	Aau72542 Arabidops
871	19	95.0	370	4	AAU00700	Aau00700 Mouse FCT	944	19	95.0	386	3	AAU72542	Aau72542 Arabidops
872	19	95.0	370	4	AAU00700	Aau00700 Mouse FCT	945	19	95.0	386	3	AAU72542	Aau72542 Arabidops
873	19	95.0	370	4	AAU00700	Aau00700 Mouse FCT	946	19	95.0	386	3	AAU72542	Aau72542 Arabidops
874	19	95.0	370	4	AAU00700	Aau00700 Mouse FCT	947	19	95.0	386	3	AAU72542	Aau72542 Arabidops
875	19	95.0	370	4	AAU00700	Aau00700 Mouse FCT	948	19	95.0	386	3	AAU72542	Aau72542 Arabidops
876	19	95.0	370	4	AAU00700	Aau00700 Mouse FCT	949	19	95.0	386	3	AAU72542	Aau72542 Arabidops
877	19	95.0	370	4	AAU00700	Aau00700 Mouse FCT	950	19	95.0	386	3	AAU72542	Aau72542 Arabidops
878	19	95.0	370	4	AAU00700	Aau00700 Mouse FCT	951	19	95.0	386	3	AAU72542	Aau72542 Arabidops
879	19	95.0	370	4	AAU00700	Aau00700 Mouse FCT	952	19	95.0	386	3	AAU72542	Aau72542 Arabidops
880	19	95.0	370	4	AAU00700	Aau00700 Mouse FCT	953	19	95.0	386	3	AAU72542	Aau72542 Arabidops
881	19	95.0	371	5	ABP29798	Abp29798 Streptoco	954	19	95.0	386	3	AAU72542	Aau72542 Arabidops
882	19	95.0	371	5	ABP29798	Abp29798 Streptoco	955	19	95.0	386	3	AAU72542	Aau72542 Arabidops
883	19	95.0	371	5	ABP29798	Abp29798 Streptoco	956	19	95.0	386	3	AAU72542	Aau72542 Arabidops
884	19	95.0	371	5	ABP29798	Abp29798 Streptoco	957	19	95.0	386	3	AAU72542	Aau72542 Arabidops
885	19	95.0	371	5	ABP29798	Abp29798 Streptoco	958	19	95.0	386	3	AAU72542	Aau72542 Arabidops
886	19	95.0	371	5	ABP29798	Abp29798 Streptoco	959	19	95.0	386	3	AAU72542	Aau72542 Arabidops
887	19	95.0	371	5	ABP29798	Abp29798 Streptoco	960	19	95.0	386	3	AAU72542	Aau72542 Arabidops
888	19	95.0	371	5	ABP29798	Abp29798 Streptoco	961	19	95.0	386	3	AAU72542	Aau72542 Arabidops
889	19	95.0	371	5	ABP29798	Abp29798 Streptoco	962	19	95.0	386	3	AAU72542	Aau72542 Arabidops
890	19	95.0	371	5	ABP29798	Abp29798 Streptoco	963	19	95.0	386	3	AAU72542	Aau72542 Arabidops
891	19	95.0	371	5	ABP29798	Abp29798 Streptoco	964	19	95.0	386	3	AAU72542	Aau72542 Arabidops
892	19	95.0	371	5	ABP29798	Abp29798 Streptoco	965	19	95.0	386	3	AAU72542	Aau72542 Arabidops
893	19	95.0	371	5	ABP29798	Abp29798 Streptoco	966	19	95.0	386	3	AAU72542	Aau72542 Arabidops
894	19	95.0	371	5	ABP29798	Abp29798 Streptoco	967	19	95.0	386	3	AAU72542	Aau72542 Arabidops
895	19	95.0	371	5	ABP29798	Abp29798 Streptoco	968	19	95.0	386	3	AAU72542	Aau72542 Arabidops
896	19	95.0	371	5	ABP29798	Abp29798 Streptoco	969	19	95.0	386	3	AAU72542	Aau72542 Arabidops
897	19	95.0	371	5	ABP29798	Abp29798 Streptoco	970	19	95.0	386	3	AAU72542	Aau72542 Arabidops
898	19	95.0	371	5	ABP29798	Abp29798 Streptoco	971	19	95.0	386	3	AAU72542	Aau72542 Arabidops
899	19	95.0	371	5	ABP29798	Abp29798 Streptoco	972	19	95.0	386	3	AAU72542	Aau72542 Arabidops
900	19	95.0	371	5	ABP29798	Abp29798 Streptoco	973	19	95.0	386	3	AAU72542	Aau72542 Arabidops
901	19	95.0	371	5	ABP29798	Abp29798 Streptoco	974	19	95.0	386	3	AAU72542	Aau72542 Arabidops

975 19 95.0 403 3 AAG50762 Arabidops  
976 19 95.0 403 3 AAG13524 Arabidops  
977 19 95.0 403 4 ABB58218 Drosophl  
978 19 95.0 404 3 AAG53518 Arabidops  
979 19 95.0 404 3 AAG39960 Arabidops  
980 19 95.0 404 6 ABU40766 Protein e  
981 19 95.0 405 2 AAY14954 Amino aci  
982 19 95.0 405 4 AAB83254 Rat FATP4  
983 19 95.0 405 6 ABM69291 Photorhab  
984 19 95.0 406 2 AAR21549 Human Try  
985 19 95.0 406 6 AAR23329 Protein e  
986 19 95.0 407 4 AAM39361 Human pol  
987 19 95.0 407 5 ABP27964 Streptoco  
988 19 95.0 407 5 ABP61471 Human NF-  
989 19 95.0 407 6 ABU46908 Protein e  
990 19 95.0 408 4 AAG91382 C Glutami  
991 19 95.0 409 5 AAE30494 Fruit fly  
992 19 95.0 410 2 AAR54663 Transcrip  
993 19 95.0 410 2 AAR89212 Transcrip  
994 19 95.0 410 5 AAE30503 Fruit fly  
995 19 95.0 411 4 AAU56508 Propionib  
996 19 95.0 411 6 ABM53027 Propionib  
997 19 95.0 411 6 ABU50525 Protein e  
998 19 95.0 412 3 AAB58972 Breast an  
999 19 95.0 412 3 AAB57091 Human pro  
1000 19 95.0 412 6 ABU14941 Protein e

## ALIGNMENT'S

RESULT 1  
AAM47151  
ID AAM47151 standard; peptide; 6 AA.  
AC AAM47151;  
XX  
DT 12-FEB-2002 (first entry)  
XX  
DE S chrysomallus actinomycin biosynthesis protein acmc fragment #9.  
DE Modular enzyme system; cyclic gene synthesis; repetitive coding sequence;  
KW antibiotic; non-ribosomal peptide synthetase; NRPS; PKS;  
KW polyketide synthase; actinomycin biosynthesis.  
XX  
OS Streptomyces chrysomallus.  
OS Synthetic.  
XX WO200181564-A2.  
XX  
PN 01-NOV-2001.  
PD  
XX 25-APR-2001; 2001WO-DE001578.  
XX  
XX 26-APR-2000; 2000DE-01021267.  
XX  
XX (ACTI-) ACTINODRUG PHARM GMBH.  
XX  
XX Schauwecker F;  
XX  
XX WPI; 2002-049276/06.  
DR N-PSDB; ABA03345.  
XX  
XX Preparing DNA encoding modular protein for e.g. producing new enzymes for  
PT synthesis of polyketide antibiotics, comprises cyclic integration of  
PT fragments into a vector.  
XX  
XX Example 3; Page 54; 83pp; German.  
XX  
XX The present invention relates to the preparation of DNA, in a circular  
CC vector, that encodes one or more segments of a modular polypeptide. DNA  
CC or DNA libraries produced this way are used to produce modular  
CC polypeptides, particularly enzymes, which can be used to act on

CC substrates to produce compounds for therapeutic testing. Enzymes of  
CC particular interest are those involved in non-ribosomal peptide synthesis  
CC or polyketide synthesis, and compounds for testing are particularly  
CC macroide antibiotics, including penicillins, vancomycins or  
CC erythromycins, but may also be modular receptors. The present sequence is  
CC a fragment of a protein encoded by a Streptomyces chrysomallus  
CC actinomycin biosynthesis gene which was used in a plasmid in the  
CC exemplification of the invention  
XX  
XX Sequence 6 AA;

SQ  
Query Match 95.0%; Score 19; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAEF 5  
Db 3 VAEF 6

RESULT 2  
AAB07871  
ID AAB07871 standard; peptide; 7 AA.

XX AAB07871;  
AC AAB07871;  
XX 14-NOV-2000 (first entry)  
XX  
XX A beta-secretase inhibitor peptide.  
DE  
XX Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;  
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;  
KW inhibitor.  
XX  
XX Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 3 /note="hydroxyethylene"  
FT

XX WO200047618-A2.

XX 17-AUG-2000.

XX 10-FEB-2000; 2000WO-US003819.

XX 10-FEB-1999; 99US-0119571P.

XX 15-JUN-1999; 99US-0139172P.

XX (ELAN-) ELAN PHARM INC.

XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;

XX Sinha S, Tateano G, Tung J, Wang S, Maconlogue L;

XX WPI; 2000-533011/48.

XX Purified beta-secretase protein used in assays to discover inhibitors

XX which can be used for the treatment of amyloidogenic diseases e.g.

XX Alzheimer's disease.

XX Disclosure; Page 12; 121pp; English.

XX The specification describes a beta-secretase enzyme. The enzyme cleaves  
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This  
CC enzyme is therefore implicated in the production of amyloid plaque  
CC components which accumulate in the brains of individuals afflicted with  
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to a  
CC mammalian subject e.g. with Alzheimer's disease or Alzheimer's disease-  
CC like pathology to test if they maintain or improve cognitive ability or  
CC reduce the plaque burden. The compounds are used for the treatment of  
CC amyloidogenic diseases e.g. Alzheimer's disease. The present sequence  
CC represents an inhibitor of beta-secretase enzyme

XX

Q Sequence 7 AA;

Query Match 95.0%; Score 19; DB 3; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
 ||||  
 b 4 VAEF 7

RESULT 3

AB07872

D AAB07872 standard; peptide; 8 AA.

X C

AAB07872;

X X

14-NOV-2000 (first entry)

X T

A beta-secretase inhibitor peptide.

X E

Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;  
 amyloid plaque component; Alzheimer's disease; amyloidogenic disease;  
 inhibitor.

X W

W

X X

S

X S

Synthetic.

X H

Key

X T

Modified-site

X 4

Location/Qualifiers

X T

/note= "hydroxyethylene"

X X

WO200047618-A2.

X N

17-AUG-2000.

X D

10-FEB-2000; 2000WO-US003819.

X F

10-FEB-1999; 99US-0119571P.

X R

15-JUN-1999; 99US-0139172P.

X R

(ELAN-) ELAN PHARM INC.

X A

Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;

X T

Sinha S, Tatsuno G, Tung J, Wang S, Mcconlogue L;

X X

WPI; 2000-533011/48.

X X

Purified beta-secretase protein used in assays to discover inhibitors

X T

which can be used for the treatment of amyloidogenic diseases e.g.

X T

Alzheimer's disease.

X X

Disclosure; Page 12; 12lpp; English.

X X

The specification describes a beta-secretase enzyme. The enzyme cleaves

X C

beta-amyloid precursor protein to produce beta-amyloid peptide. This

X C

enzyme is therefore implicated in the production of amyloid plaque

X C

components which accumulate in the brains of individuals afflicted with

X C

Alzheimer's disease. Inhibitors of beta-secretase are administered to a

X C

mammalian subject e.g. with Alzheimer's disease or Alzheimer's disease-

X C

like pathology to test if they maintain or improve cognitive ability or

X C

reduce the plaque burden. The compounds are used for the treatment of

X C

amyloidogenic diseases e.g. Alzheimer's disease. The present sequence

X C

represents an inhibitor of beta-secretase enzyme

X X

Sequence 8 AA;

Q

Query Match 95.0%; Score 19; DB 3; Length 8;

X

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

X

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

X

Y 2 VAEF 5

X

||||

X

5 VAEF 8

X

DT 03-JAN-2002 (first entry)

RESULT 4

ABB56241

ID ABB56241 standard; peptide; 8 AA.

XX AC

ABB56241;

XX AC

15-FEB-2002 (first entry)

XX XX

Vascular dementia-associated protein isoform (VPI) 441.

XX DE

Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;

XX KW

diagnosis; prognosis; gene therapy.

XX XW

Homo sapiens.

XX OS

WO200169261-A2.

XX PN

20-SEP-2001.

XX PD

14-MAR-2001; 2001WO-GB001106.

XX PF

15-MAR-2000; 2000GB-00006285.

XX PR

24-NOV-2000; 2000GB-00028734.

XX PR

28-NOV-2000; 2000US-00724391.

XX XX

(OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX PA

Herath HWAC, Parekh RB, Rohlf C;

XX PI

WPI; 2001-557937/62.

XX DR

Screening, diagnosis or prognosis of vascular dementia (VD), useful for

XX XX

determining stage of VD and monitoring the effect of VD therapy.

XX PT

comprises analyzing body fluid by 2-dimensional electrophoresis for

XX PT

features correlated with VD.

XX XX

Claim 6; Page 39; 15lpp; English.

XX XX

The invention relates to screening, diagnosis or prognosis of Vascular

XX CC

Dementia (VD) in a subject comprising analysing body fluid from the

XX CC

subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of

XX CC

features containing at least one chosen feature whose relative abundance

XX CC

correlates with the presence, absence, stage or severity of VD or

XX CC

predicts the onset or course of VD, especially detecting in a sample of

XX CC

cerebrospinal fluid (CSF) from the subject one of 223 VD-associated

XX CC

protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the

XX CC

specification. Detecting VD-associated features and VPI is useful for the

XX CC

screening, diagnosis or prognosis of VD, for determining the stage or

XX CC

severity of VD, for identifying a subject at risk of VD or for monitoring

XX CC

the effect of therapy administered to a subject having VD. Nucleic acids

XX CC

encoding a VPI or inhibiting the function of a VPI are useful for the

XX CC

treatment of VD and for gene therapy

XX XX

Sequence 8 AA;

SQ

Query Match 95.0%; Score 19; DB 4; Length 8;

XX

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

XX

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

2 VAEF 5

XX

||||

XX

2 VAEF 5

XX

DT 03-JAN-2002 (first entry)

RESULT 5

AAU28720

ID AAU28720 standard; peptide; 8 AA.

XX XX

AAU28720;

XX AC

03-JAN-2002 (first entry)

XX DT



XX DE DPI tryptic digest peptide #317.  
 XX KW Human; depression associated protein isoform; tryptic digest peptide;  
 XX KW DPI; cerebrospinal fluid; CSF; BAD; bipolar affective disorder;  
 XX KW neuropsychiatric disorder; bipolar mood disorder; neuroleptic;  
 XX KW maniac-depressive illness; schizoaffective disorder.  
 XX OS Homo sapiens.  
 XX PN WO200162787-A1.  
 XX PD 30-AUG-2001.  
 XX PF 23-FEB-2001; 2001WO-GB000786.  
 XX PR 24-FEB-2000; 2000GB-00004412.  
 XX PR 08-DEC-2000; 2000GB-00030050.  
 XX PR 12-DEC-2000; 2000US-0254830P.  
 XX PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 XX PI Herath HMAC, Parekh RB, Rohlf C, Terrett JA, Tyson KL;  
 XX PF WPI; 2001-570626/64.  
 XX PT Novel nucleic acid encoding a protein associated with bipolar affective  
 XX PT disorder, which is used for diagnosis, prophylaxis and therapy of  
 XX PT neuropsychiatric disorders, such as bipolar affective disorder.  
 XX PS Disclosure; Page 37; 153pp; English.  
 XX CC The present invention relates to the identification of depression  
 XX CC associated protein isoforms (DPIs), particularly the tryptic digest  
 XX CC peptides of these proteins. Some of the DPIs (AAU28404-AAU28625)  
 XX CC described are decreased in the cerebrospinal fluid (CSF) of BAD (bipolar  
 XX CC affective disorder) subjects, whilst other DPIs (AAU28626-AAU28887) are  
 XX CC increased in BAD subjects. Also described are peptide sequences  
 XX CC identified from DPI-45 and DPI-213 and the nucleic acid sequences they are  
 XX CC encoded by. The sequences of the invention are useful for clinical  
 XX CC screening, diagnosis, prognosis, therapy and prophylaxis of  
 XX CC neuropsychiatric disorders e.g. BAD (also known as bipolar mood disorder,  
 XX CC BP), maniac-depressive illnesses, attention deficit disorders,  
 XX CC schizoaffective disorders, and unipolar affective disorders. The present  
 XX CC sequence represents one of the DPI tryptic digest peptides of the present  
 XX CC invention  
 XX SQ Sequence 8 AA;  
 XX Query Match 95.0%; Score 19; DB 4; Length 8;  
 XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 DB |||||  
 2 VAEF 5  
 RESULT 6  
 AAU25114  
 ID AAU25114 standard; peptide; 8 AA.  
 XX AC AAU25114;  
 XX DT 18-DEC-2001 (first entry)  
 XX DE Schizophrenia-Associated Protein Isoform (SPI) peptide #343.  
 XX KW Schizophrenia-associated protein isoform; SPI; SPI-206; SPI-238; SPI-240;  
 XX KW neuroleptic; gene therapy; cerebrospinal fluid; serum; plasma.  
 XX OS Homo sapiens.  
 XX PN WO200162787-A2.  
 XX PD 30-AUG-2001.  
 XX PF 23-FEB-2001; 2001WO-GB000791.  
 XX PR 24-FEB-2000; 2000GB-00004412.  
 XX PR 08-DEC-2000; 2000GB-00030050.  
 XX PR 12-DEC-2000; 2000US-0254830P.  
 XX PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 XX PI Herath HMAC, Parekh RB, Rohlf C, Terrett JA, Tyson KL;  
 XX PF WPI; 2001-570624/64.  
 XX PT New schizophrenia associated protein isoforms and encoding nucleic acid  
 XX PT molecules, useful for treatment, diagnosis and prognosis of schizophrenia  
 XX PT and screening for potential drugs for treatment and new drug targets.  
 XX PS Disclosure; Page 36; 148pp; English.  
 XX CC The sequence represents a schizophrenia-associated protein isoform (SPI).  
 XX CC These protein isoforms, e.g. SPI-206, SPI-238 and SPI-240 are detectable  
 XX CC in cerebrospinal fluid, serum or plasma and are useful markers of  
 XX CC schizophrenia. The sequences can be used for treatment and diagnosis of  
 XX CC schizophrenia, screening, prognosis, monitoring the results of therapy,  
 XX CC identifying patients most likely to respond to a particular therapy and  
 XX CC identification of new targets for drug treatment. SPI DNA is useful as a  
 XX CC nucleic acid probe to detect the presence of nucleic acids or SPIs  
 XX SQ Sequence 8 AA;  
 XX Query Match 95.0%; Score 19; DB 4; Length 8;  
 XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 DB |||||  
 2 VAEF 5  
 RESULT 7  
 AAU26368  
 ID AAU26368 standard; peptide; 8 AA.  
 XX AC AAU26368;  
 XX DT 18-DEC-2001 (first entry)  
 XX DE Depression-Associated Protein isoform DPI-124.  
 XX KW Human; Bipolar Affective Disorder; BAD; Depression-Associated feature;  
 XX KW DP; Depression-Associated protein isoform; DPI; Cerebro-spinal fluid;  
 XX KW CSF; antidepressant; antimanic; nootropic; tranquiliser; neuroleptic;  
 XX KW attention deficient disorder; schizoaffective disorder;  
 XX KW unipolar affective disorder.  
 XX OS Homo sapiens.  
 XX PN WO200163294-A2.  
 XX PD 30-AUG-2001.  
 XX PF 23-FEB-2001; 2001WO-GB000791.  
 XX PR 24-FEB-2000; 2000GB-00004412.  
 XX PR 08-DEC-2000; 2000GB-00030050.  
 XX PR 12-DEC-2000; 2000US-0254830P.  
 XX PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 XX PI Herath HMAC, Parekh RB, Rohlf C, Terrett JA, Tyson KL;  
 XX PF WPI; 2001-570624/64.  
 XX PT New schizophrenia associated protein isoforms and encoding nucleic acid  
 XX PT molecules, useful for treatment, diagnosis and prognosis of schizophrenia  
 XX PT and screening for potential drugs for treatment and new drug targets.  
 XX PS Disclosure; Page 36; 148pp; English.  
 XX CC The sequence represents a schizophrenia-associated protein isoform (SPI).  
 XX CC These protein isoforms, e.g. SPI-206, SPI-238 and SPI-240 are detectable  
 XX CC in cerebrospinal fluid, serum or plasma and are useful markers of  
 XX CC schizophrenia. The sequences can be used for treatment and diagnosis of  
 XX CC schizophrenia, screening, prognosis, monitoring the results of therapy,  
 XX CC identifying patients most likely to respond to a particular therapy and  
 XX CC identification of new targets for drug treatment. SPI DNA is useful as a  
 XX CC nucleic acid probe to detect the presence of nucleic acids or SPIs  
 XX SQ Sequence 8 AA;  
 XX Query Match 95.0%; Score 19; DB 4; Length 8;  
 XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 DB |||||  
 2 VAEF 5

WPI; 2001-502081/65.  
Preparation for diagnosing or treating bipolar affected disorder (BAD) or unipolar depression, or for screening for modulators, comprises a BAD-associated protein isoform.  
Claim 8; Page 37; 163pp; English.  
The invention relates to a preparation comprising an isolated Bipolar Affected Disorder (BAD)-Associated Protein Isoform (DPIs). The DPI's are used to screen, diagnose or prognosis of BAD or unipolar depression, determine the stage or severity of BAD or unipolar depression, identify a subject at risk of developing BAD or unipolar depression, or monitor the effect of therapy in a subject. They are also used to screen for or identify agents that interact with a DPI. These agents, antibodies against the DPIs, and nucleic acids encoding the DPIs are used to treat or prevent BAD or unipolar depression. Diseases that can be treated are attention deficient disorder, a schizoaffective disorder, a bipolar or a unipolar affective disorder. The DPIs are used in proteomics. The proteomic approach of using DPIs for screening, diagnosis or prognosis of BAD or unipolar depression overcomes the problems of using gene expression analysis, such as not being able to obtain central nervous system (CNS) tissue from a living patient under normal circumstances. The present sequence is a DIP increased in the CSF (cerebro-spinal fluid) of subjects having BAD

Sequence 8 AA;  
Query Match 95.0%; Score 19; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||||  
2 VAEF 5

RESULT 8  
AAU15458  
ID AAU15458 standard; peptide; 8 AA.  
AC AAU15458;  
24-OCT-2001 (first entry)  
Schizophrenia-associated isoform peptide #343.  
Schizophrenia; neuroleptic; diagnostic; neuropsychiatric disorder; neurological disorder; neuropathy.  
Homo sapiens.  
WO200163293-A2.  
30-AUG-2001.  
23-FEB-2001; 2001WO-GB000783.  
24-FEB-2000; 2000GB-00004415.  
28-DEC-2000; 2000US-00750395.  
(OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
Herath HMCAC, Parekh RB, Rohlf C;  
WPI; 2001-502868/55.  
Diagnosing and monitoring Schizophrenia by detecting the presence of Schizophrenia Associated Features and Schizophrenia Associated Protein Isoforms in samples of cerebrospinal fluid.  
Claim 6; Page 36; 160pp; English.

The invention relates to methods and compositions for screening, diagnosis and prognosis of Schizophrenia. The method involves detecting the presence of Schizophrenia (SCH) Associated Features (SFS) and SCH Associated Protein Isoforms (SPIs) in samples, e.g. by electrophoresis, immunosay or hybridisation assay, for diagnosing and monitoring SCH, studying the effectiveness of treatments and for identifying potential therapeutic agents. The method is used for (1) screening or diagnosis of SCH and the relative abundance of at least 1 chosen feature correlates with the presence or absence of SCH; and (2) monitoring the effect of therapy administered to a subject with SCH and the relative abundance of at least 1 chosen feature which correlates with the severity of SCH. The expression and activity of the SFS, SPIs and related molecules (e.g. secondary messengers) are studied to diagnose SCH, monitor the progress of the disorder and the effectiveness of treatment and as targets to identify and produce potential therapeutic agents for the treatment of SCH. The paucity of detectable neurologic defects distinguishes neuropsychiatric disorders such as SCH from neurological disorders, where manifestations of anatomical and biochemical changes have been identified in many cases. Consequently the identification and characterisation of cellular and/or molecular causative defects and neuropathies are necessary for improved treatment of neuropsychiatric disorders. AAU1514-CC AAU15762 represent the amino acid sequences of schizophrenia-associated isoforms used in the method of the invention

Sequence 8 AA;  
Query Match 95.0%; Score 19; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||||  
2 VAEF 5

RESULT 9  
AAR78909  
ID AAR78909 standard; peptide; 9 AA.  
AC AAR78909;  
25-MAR-2003 (revised)  
27-MAR-1996 (first entry)  
MAGE 3 105-113 cytotoxic T lymphocyte epitope.  
MAGE 3 105-113; cytotoxic T; CTL; epitope; helper T; HTL; cell; lymphocyte; antigens; treatment; disease prevention; tumours; cancer; melanomas.  
Homo sapiens.  
WO9522317-A1.  
24-AUG-1995.  
16-FEB-1995; 95WO-US002121.  
16-FEB-1994; 94US-00197484.  
(CYTE-) CYTEL CORP.  
Vitiello MA, Chesnut RW, Sette AD, Celis E, Grey H;  
WPI; 1995-302545/39.  
Compn. inducing cytotoxic T lymphocyte response to pref. viral, bacterial, parasitic or tumour antigens - useful in the treatment and prevention of diseases associated with the antigen e.g. hepatitis B.  
Example 13; Page 71; 109pp; English.  
A compsn. which induces a cytotoxic T lymphocyte (CTL) response to a

CC human MAGE antigen (Ag) in a mammal comprises, a MAGE CTL Ag response  
 CC inducing peptide (i.e. AAR78904 to AAR78917) and a lipid conjugated  
 CC helper T cell inducing peptide. The compsn. is useful in the treatment  
 CC and prevention of MAGE tumour Ag associated diseases, e.g. melanoma  
 CC cancers. (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 9 AA;  
 Query Match 95.0%; Score 19; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 DB 2 VAEF 5  
 RESULT 10  
 ABB77871  
 ID ABB77871 standard; peptide; 9 AA.  
 XX  
 AC ABB77871;  
 XX  
 DT 27-SEP-2002 (first entry)  
 XX  
 DE A beta-amyloid precursor protein (APP) inhibitor peptide.  
 XX  
 KW Beta-amyloid precursor protein; APP; APP inhibitor peptide; BACE;  
 KW beta site APP cleaving enzyme; protein coordinate data; APP751;  
 KW Swedish family mutation; Alzheimer's disease.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 5 /label= Sta  
 FT FT /note= "statine"  
 XX  
 PN WO200225276-A1.  
 XX  
 PD 28-MAR-2002.  
 XX  
 XX 19-SEP-2001; 2001WO-US029387.  
 PF  
 XX 22-SEP-2000; 2000US-0234576P.  
 PR  
 XX (AMHP ) AMERICAN HOME PROD CORP.  
 PA  
 XX Choppa R, Svenson K, Annis B, Akopian TN, Bard J, Stahl ML;  
 PI Somers WS;  
 PI WPI; 2002-519081/55.  
 DR  
 XX Crystallized complex of beta-site amyloid precursor protein (APP)  
 PT cleaving enzyme (BACE) and APP inhibitor peptide useful for identifying  
 PT agents that interact with active site of BACE or active site of APP  
 PT binding protein or peptide.  
 XX  
 PS Claim 1; Page 22; 88pp; English.  
 XX  
 CC The present sequence represents a beta-amyloid precursor protein (APP)  
 CC inhibitor peptide. This peptide inhibits binding between APP and BACE  
 CC (beta site APP cleaving enzyme). The specification describes a  
 CC crystallized complex of BACE and the present APP inhibitor. Protein  
 CC coordinate data for BACE is given in the specification. The APP inhibitor  
 CC peptide is based on the P10 to P4' APP751 Swedish family mutation. The  
 CC crystallized complex is used for identifying an agent that interacts with  
 CC an active site of BACE or an active site of an APP binding protein or  
 CC peptide. The agents are useful in the treatment and/or prevention of  
 CC Alzheimer's disease  
 XX  
 SQ Sequence 9 AA;

Query Match 95.0%; Score 19; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 DB 6 VAEF 9  
 RESULT 11  
 ABB06593  
 ID ABB06593 standard; peptide; 10 AA.  
 XX  
 AC ABB06593;  
 XX  
 DT 31-MAY-2002 (first entry)  
 XX  
 DE Beta-secretase related peptide SEQ ID NO:197.  
 XX  
 KW Beta-secretase; enzyme; cleavage site; amyloid protein precursor; APP;  
 KW aspartyl protease; neuroprotective; nootropic; beta-secretase inhibitor;  
 KW Alzheimer's disease.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200206306-A2.  
 PD 24-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US023035.  
 XX  
 PR 19-JUL-2000; 2000US-0219795P.  
 PR 12-MAR-2001; 2001US-0275251P.  
 XX  
 PA (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 PI Yan R, Tomasselli AG, Gurney ME, Emmons TL, Bienkowski MJ;  
 PI Heinrichson RL;  
 XX  
 DR WPI; 2002-216995/27.  
 XX  
 PT Novel substrates for human aspartyl protease useful for identifying  
 PT modulators of beta secretase activity of aspartyl protease for treating  
 PT Alzheimer's disease.  
 XX  
 PS Disclosure; Page 188; 188pp; English.  
 XX  
 CC The present invention describes an isolated peptide (I) comprising a  
 CC sequence of at least four amino acids, where the peptide is a substrate  
 CC for conducting aspartyl protease assays. (I) has neuroprotective and  
 CC nootropic activities, and can be used as an inhibitor of beta-secretase  
 CC activity. A beta-secretase modulator from the present invention can be  
 CC used for inhibiting beta-secretase activity in vivo, and in the  
 CC manufacture of a medicament for the treatment of Alzheimer's disease.  
 CC Pharmaceutical compositions from the present invention can be used for  
 CC treating a disease or condition characterised by an abnormal beta-  
 CC secretase activity. (I) is useful for identifying agents that modulate  
 CC the activity of human Asp2 aspartyl protease (Hu-Asp2). (I) is useful as  
 CC a core structure to construct derivatives. ABL49914 to ABL49925 and  
 CC ABB06409 to ABB06593 represent sequences used in the exemplification of  
 CC the present invention  
 XX  
 SQ Sequence 10 AA;  
 Query Match 95.0%; Score 19; DB 5; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 DB 5 VAEF 8

RESULT 12  
BB06592  
D ABB06592 standard; peptide; 12 AA.  
X C ABB06592;  
X T 31-MAY-2002 (first entry)  
X E Beta-secretase related peptide SEQ ID NO:196.  
X Beta-secretase; enzyme; cleavage site; amyloid protein precursor; APP;  
W aspartyl protease; neuroprotective; nontropic; beta-secretase inhibitor;  
W Alzheimer's disease.  
X X Homo sapiens.  
X S Synthetic.  
X W WO200206306-A2.  
X N 24-JAN-2002.  
X D 19-JUL-2001; 2001WO-US023035.  
X F 19-JUL-2000; 2000US-0219795P.  
X R 12-MAR-2001; 2001US-0275251P.  
X R (PHAA ) PHARMACIA & UPJOHN CO.  
X A Yan R. Tomasselli AG, Gurney ME, Emmons TL, Bienkowski MJ;  
X I Heinrichson RH;  
X WPI; 2002-216995/27.  
X Novel substrates for human aspartyl protease useful for identifying  
T modulators of beta secretase activity of aspartyl protease for treating  
T Alzheimer's disease.  
X S Disclosure; Page 188; 189pp; English.  
X The present invention describes an isolated peptide (I) comprising a  
C sequence of at least four amino acids, where the peptide is a substrate  
C for conducting aspartyl protease assays. (I) has neuroprotective and  
C nontropic activities, and can be used as an inhibitor of beta-secretase  
C activity. A beta-secretase modulator from the present invention can be  
C used for inhibiting beta-secretase activity in vivo, and in the  
C manufacture of a medicament for the treatment of Alzheimer's disease.  
C Pharmaceutical compositions from the present invention can be used for  
C treating a disease or condition characterised by an abnormal beta-  
C secretase activity. (I) is useful for identifying agents that modulate  
C the activity of human Asp2 aspartyl protease (Hu-Asp2). (I) is useful as  
C a core structure to construct derivatives. ABL49914 to ABL49925 and  
C ABB06409 to ABB06593 represent sequences used in the exemplification of  
C the present invention  
X Q Sequence 12 AA;  
Query Match 95.0%; Score 19; DB 5; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Y 2 VAEF 5  
b 5 VAEF 8  
|||  
|||  
RESULT 13  
BG78404  
D ABB06592 standard; peptide; 12 AA.  
X C ABB06592;  
X T 31-MAY-2002 (first entry)  
X E Beta-secretase related peptide SEQ ID NO:196.  
X Beta-secretase; enzyme; cleavage site; amyloid protein precursor; APP;  
W aspartyl protease; neuroprotective; nontropic; beta-secretase inhibitor;  
W Alzheimer's disease.  
X X Homo sapiens.  
X S Synthetic.  
X W WO200206306-A2.  
X N 24-JAN-2002.  
X D 19-JUL-2001; 2001WO-US023035.  
X F 19-JUL-2000; 2000US-0219795P.  
X R 12-MAR-2001; 2001US-0275251P.  
X R (PHAA ) PHARMACIA & UPJOHN CO.  
X A Yan R. Tomasselli AG, Gurney ME, Emmons TL, Bienkowski MJ;  
X I Heinrichson RH;  
X WPI; 2002-216995/27.  
X Novel substrates for human aspartyl protease useful for identifying  
T modulators of beta secretase activity of aspartyl protease for treating  
T Alzheimer's disease.  
X S Disclosure; Page 188; 189pp; English.  
X The present invention describes an isolated peptide (I) comprising a  
C sequence of at least four amino acids, where the peptide is a substrate  
C for conducting aspartyl protease assays. (I) has neuroprotective and  
C nontropic activities, and can be used as an inhibitor of beta-secretase  
C activity. A beta-secretase modulator from the present invention can be  
C used for inhibiting beta-secretase activity in vivo, and in the  
C manufacture of a medicament for the treatment of Alzheimer's disease.  
C Pharmaceutical compositions from the present invention can be used for  
C treating a disease or condition characterised by an abnormal beta-  
C secretase activity. (I) is useful for identifying agents that modulate  
C the activity of human Asp2 aspartyl protease (Hu-Asp2). (I) is useful as  
C a core structure to construct derivatives. ABL49914 to ABL49925 and  
C ABB06409 to ABB06593 represent sequences used in the exemplification of  
C the present invention  
X Q Sequence 12 AA;  
Query Match 95.0%; Score 19; DB 5; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Y 2 VAEF 5  
b 5 VAEF 8  
|||  
|||

DT 15-NOV-2002 (first entry)  
XX Memapsin 2 substrate specificity determination peptide #10.  
XX Human; memapsin 2; aspartic protease; beta secretase;  
XX degenerative disease; Alzheimer's disease; amyloid precursor protein;  
XX APP; neuroprotective; nontropic; inhibitor;  
XX substrate side-chain preference.  
XX Homo sapiens.  
XX Synthetic.  
XX WO200253594-A2.  
XX 11-JUL-2002.  
XX 28-DEC-2001; 2001WO-US050826.  
XX 28-DEC-2000; 2000US-0258705P.  
XX 14-MAR-2001; 2001US-0275756P.  
XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.  
XX (UNII ) UNIV ILLINOIS FOUND.  
XX Tang JJN, Koelsch G, Ghosh AK;  
XX WPI; 2002-619088/66.  
XX New memapsin 2 activity inhibitor useful in treatment of e.g. Alzheimer's  
XX disease.  
XX Example 5; Page 51; 74pp; English.  
XX The invention relates to an inhibitor of catalytically active memapsin 2  
XX (an aspartic protease which can cleave at beta secretase sites), which  
XX binds to the active site of memapsin 2 defined by the presence of two  
XX catalytic aspartic residues and substrate binding cleft. Also included is  
XX a method of determination of the substrate side-chain preference in  
XX memapsin 2 sub-sites comprising: (a) reacting a mixture of memapsin 2  
XX substrates with memapsin 2, and determining the sub-site preference of the  
XX mixture of memapsin 2 by determining relative initial hydrolysis rates of the  
XX library of memapsin 2 inhibitors containing a base sequence taken from  
XX OX99-2 (Glu-Val-An-Leu-Ala-Ala-Glu-Phe), probing the library of  
XX inhibitors with memapsin 2 which binds to several inhibitors to generate  
XX several bound memapsin 2, and detecting the bound memapsin 2 with an  
XX antibody raised to memapsin 2, and an alkaline phosphatase conjugated  
XX secondary antibody. The inhibitors may be used in the manufacture of a  
XX medicament for the treatment of Alzheimer's disease since memapsin 2 may  
XX be involved in the cleavage of amyloid precursor protein (APP), and for  
XX determining the substrate side-chain preference in memapsin 2 sub-sites.  
XX The present sequence represents a sub-site variant peptide used to  
XX determine the substrate specificity of human memapsin 2  
XX Sequence 12 AA;  
Query Match 95.0%; Score 19; DB 5; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 VAEF 5  
Db 5 VAEF 8  
|||  
|||  
RESULT 14  
ADD35467  
ID ADD35467 standard; peptide; 12 AA.  
XX ADD35467;  
XX ADD35467;  
XX 15-JAN-2004 (first entry)  
DT DT

DE Escherichia coli Dnak peptide #3.  
 XX crystallised recombinant protein; metabolism; Staphylococcus aureus;  
 KW Streptococcus pneumoniae; Helicobacter pylori; Escherichia coli;  
 KW Pseudomonas aeruginosa; vaccine.  
 XX  
 OS Escherichia coli.  
 XX  
 PN WO2003044185-A2.  
 XX  
 PD 30-MAY-2003.  
 XX  
 XX 21-NOV-2002; 2002WO-CA001768.  
 PF  
 XX 21-NOV-2001; 2001US-0332160P.  
 PR  
 PR 27-NOV-2001; 2001US-0333661P.  
 PR  
 PR 27-NOV-2001; 2001US-0333665P.  
 PR  
 PR 18-DEC-2001; 2001US-0341770P.  
 PR  
 PR 19-DEC-2001; 2001US-0341954P.  
 PR  
 PR 19-DEC-2001; 2001US-0342003P.  
 PR  
 PR 20-DEC-2001; 2001US-0342542P.  
 PR  
 PR 21-DEC-2001; 2001US-0344252P.  
 PR  
 PR 28-DEC-2001; 2001US-0343570P.  
 PR  
 PR 28-DEC-2001; 2001US-0343606P.  
 PR  
 PR 28-DEC-2001; 2001US-0343679P.  
 XX  
 PA (AFPI-) AFFINIUM PHARM INC.  
 XX  
 XX Edwards A, Dharamsi A, Vedadi M, Alam MZ, Awrey D, Beattie B;  
 PI Canadian V, Domagala M, Houston S, Mansoury K, Necakov S, Nethery K;  
 PI Ng I, Pinder B, Sheldrick B, Vallee F, Wrezel O;  
 XX  
 XX WPI; 2003-513596/48.  
 DR  
 XX New crystallized recombinant polypeptides from Staphylococcus aureus,  
 PT Streptococcus pneumoniae, Helicobacter pylori or Pseudomonas aeruginosa  
 PT involved in general metabolism, useful as drug targets for pathogenic  
 PT bacteria.  
 XX  
 PS Disclosure; SEQ ID NO 66; 277pp; English.  
 XX  
 XX The invention comprises a crystallised recombinant protein that is  
 CC involved in general metabolism, the recombinant protein may be from  
 CC Staphylococcus aureus, Streptococcus pneumoniae, Helicobacter pylori,  
 CC Escherichia coli or Pseudomonas aeruginosa. The crystallised recombinant  
 CC protein of the invention is useful in the prevention (vaccine) or  
 CC treatment of a disease or disorder caused by S. pneumoniae, H. pylori, E.  
 CC coli or P. aeruginosa. The present amino acid sequence was used in the  
 CC exemplification of the invention.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 95.0%; Score 19; DB 7; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 Db 9 VAEF 12  
 RESULT 15  
 AAB07889  
 ID AAB07889 standard; peptide; 13 AA.  
 XX  
 AC AAB07889;  
 XX  
 XX 14-NOV-2000 (first entry)  
 DT  
 XX A beta-secretase inhibitor peptide.  
 DE  
 XX Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;  
 KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;  
 XX

KW inhibitor.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200047618-A2.  
 XX  
 XX 17-AUG-2000.  
 XX  
 XX 10-FEB-2000; 2000WO-US003819.  
 XX  
 PR 10-FEB-1999; 99US-0119571P.  
 PR 15-JUN-1999; 99US-0139172P.  
 XX  
 PA (ELAN-) ELAN PHARM INC.  
 XX  
 XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;  
 PI Sinha S, Tatsuno G, Tung J, Wang S, Mcconlogue L;  
 XX  
 XX WPI; 2000-533011/48.  
 DR  
 XX Purified beta-secretase protein used in assays to discover inhibitors  
 PT which can be used for the treatment of amyloidogenic diseases e.g.  
 PT Alzheimer's disease.  
 XX  
 PS Claim 33; Page 24; 121pp; English.  
 XX  
 CC The specification describes a beta-secretase enzyme. The enzyme cleaves  
 CC beta-amyloid precursor protein to produce beta-amyloid peptide. This  
 CC enzyme is therefore implicated in the production of amyloid plaque  
 CC components which accumulate in the brains of individuals afflicted with  
 CC Alzheimer's disease. Inhibitors of beta-secretase are administered to a  
 CC mammalian subject e.g. with Alzheimer's disease or Alzheimer's disease-  
 CC like pathology to test if they maintain or improve cognitive ability or  
 CC reduce the plaque burden. The compounds are used for the treatment of  
 CC amyloidogenic diseases e.g. Alzheimer's disease. The present sequence  
 CC represents an inhibitor of beta-secretase enzyme  
 XX  
 SQ Sequence 13 AA;  
 Query Match 95.0%; Score 19; DB 3; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 Db 10 VAEF 13  
 RESULT 16  
 AAM99276  
 ID AAM99276 standard; peptide; 13 AA.  
 XX  
 XX AAM99276;  
 AC  
 XX 07-DEC-2001 (first entry)  
 DT  
 XX Vaccine related MHC ligand peptide SEQ ID NO:379.  
 DE  
 XX Glutamic acid; glutamine; vaccine; major histocompatibility complex; MHC;  
 KW immunomodulator; antiallergic; endocrine; neuroprotectant; virucidal,  
 KW bactericidal; antiparasitic; fungicidal; cytostatic; medicine;  
 KW pharmaceutical; immune disorder; immune deficiency; autoimmune;  
 KW hypersensitivity; allergy; graft rejection; infection; hormonal disorder;  
 KW central nervous system disease; cancer; melanoma; anti-melanoma vaccine;  
 KW human immunodeficiency virus.  
 XX  
 XX Mycobacterium leprae.  
 OS  
 XX WO200170772-A2.  
 PN  
 XX 27-SEP-2001.  
 PD  
 XX 22-MAR-2001; 2001WO-FR000872.  
 PF

X 23-MAR-2000; 2000FR-00003711.  
 X (FABR ) FABRE MEDICAMENT SA PIERRE.  
 X Klinguer-Hamour C, Corvaia N, Beck A, Goetsch L;  
 X WPI; 2001-611470/70.  
 X Stabilized pharmaceutical containing N-terminal glutamic acid or  
 T glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt  
 T with strong acid.  
 X Claim 9; Page 96; 149pp; French.  
 X The present invention describes a pharmaceutical compound (I) that  
 C contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue in  
 C the form of an addition salt with a strong, physiologically acceptable  
 C acid (II). Also described are: (a) a pharmaceutical composition  
 C containing at least one (I); (b) a vaccine containing at least one (I)  
 C where this is a major histocompatibility complex (MHC) ligand (Ia); (c) a  
 C method for in vitro diagnosis of diseases associated with the presence of  
 C (Ia); (d) a kit for method (c) that includes a (Ia); and (e) a process  
 C for preparing (I). (I) has immunomodulator, endocrine, antiallergic,  
 C neuroprotectant, virucidal, bactericidal, antiparasitic, fungicidal and  
 C cytostatic activities. (I) are useful, in human or veterinary medicine,  
 C in pharmaceutical compositions (for treating immune disorders, e.g.  
 C immune deficiency, autoimmune states, hypersensitivity, allergy, graft  
 C rejection, infection, hormonal disorders and central nervous system  
 C diseases), also, where (I) is a MHC ligand (Ia), in vaccines for  
 C treatment or prevention of: (i) viral, bacterial, parasitic or fungal  
 C infections; or (ii) of cancers. A particular application is in anti-  
 C melanoma vaccines. (I) are also useful for in vitro diagnosis of diseases  
 C associated with interactions between MHC and (I), e.g. melanoma and human  
 C immunodeficiency virus infection. AM98898 to AM99592 represent peptides  
 C which can be used in pharmaceutical compounds from the present invention  
 X Q Sequence 13 AA;  
 Query Match 95.0%; Score 19; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Y 2 VAEF 5  
 b 4 VAEF 7  
 RESULT 17  
 AB07888  
 D AAB07888 standard; peptide; 14 AA.  
 C AAB07888;  
 X 14-NOV-2000 (first entry)  
 T A peptide fragment derived from beta-amyloid precursor protein.  
 X Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;  
 W amyloid plaque component; Alzheimer's disease; amyloidogenic disease;  
 W inhibitor.  
 X Homo sapiens.  
 S WO200047618-A2.  
 N 17-AUG-2000.  
 X 10-FEB-2000; 2000WO-US003819.  
 F 10-FEB-1999; 99US-0119571P.  
 R 15-JUN-1999; 99US-0139172P.  
 X

PA (ELAN-) ELAN PHARM INC.  
 XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;  
 PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;  
 XX WPI; 2000-533011/48.  
 DR Purified beta-secretase protein used in assays to discover inhibitors  
 XX which can be used for the treatment of amyloidogenic diseases e.g.  
 PT Alzheimer's disease.  
 XX Disclosure; Page 12; 121pp; English.  
 XX The specification describes a beta-secretase enzyme. The enzyme cleaves  
 CC beta-amyloid precursor protein to produce beta-amyloid peptide. This  
 CC enzyme is therefore implicated in the production of amyloid plaque  
 CC components which accumulate in the brains of individuals afflicted with  
 CC Alzheimer's disease. Inhibitors of beta-secretase are administered to a  
 CC mammalian subject e.g. with Alzheimer's disease or Alzheimer's disease-  
 CC like pathology to test if they maintain or improve cognitive ability or  
 CC reduce the plaque burden. The compounds are used for the treatment of  
 CC amyloidogenic diseases e.g. Alzheimer's disease. The present sequence  
 CC represents a peptide derived from beta-amyloid precursor protein  
 XX Sequence 14 AA;  
 SQ Query Match 95.0%; Score 19; DB 3; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 Db 11 VAEF 14  
 RESULT 18  
 AA097460  
 ID RAM97460 standard; peptide; 14 AA.  
 XX AC RAM97460;  
 XX 24-JAN-2002 (first entry)  
 DT Human peptide #735 encoded by a SNP oligonucleotide.  
 XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;  
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
 KW multifactorial disease; autoimmune disease; infection;  
 KW nervous system disease.  
 XX Homo sapiens.  
 OS WO200147944-A2.  
 PN 05-JUL-2001.  
 XX 28-DEC-2000; 2000WO-US035498.  
 PF 28-DEC-1999; 99US-0173419P.  
 PR 27-DEC-2000; 2000US-00173419.  
 XX (CURA-) CURAGEN CORP.  
 XX Shimkets RA, Leach M;  
 XX WPI; 2001-465210/50.  
 DR Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
 PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,  
 PT



Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
b 4 VAEF 7

RESULT 21  
AAW42943  
D AAW42943 standard; peptide; 20 AA.  
X  
C AAW42943;  
X  
T 28-APR-1998 (first entry)  
X  
E Immunogenic Hepatitis A virus peptide YK-1368.  
X  
W Immunogenic peptide; immunogenic epitope; P3A protein; immune response;  
X antibody.  
X Synthetic.  
X Hepatitis A virus.  
X WO9740147-A1.  
X  
D 30-OCT-1997.  
X  
F 18-APR-1997; 97WO-US006891.  
X  
R 19-APR-1996; 96US-0015644P.  
X  
A (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
X  
I Fields HA, Khudyakov YE;  
X WPI; 1997-535831/49.  
X  
T Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune  
T response to HAV in a mammal or to detect the presence of antibodies  
T against HAV in a mammal.  
X  
S Claim 33; Page 115; 140pp; English.  
X  
C Peptides AAW42943-46 are immunogenic peptides corresponding to  
C immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are  
C substantially similar to a portion of the amino acid sequence of the P3A  
C protein of HAV corresponding to amino acids 1423-1496. The present  
C peptide is derived from amino acids 1421-1440 and has a reactivity of  
C 63.4% with acute sera. Compositions containing the peptides can be used  
C to induce an immune response to HAV in a mammal. The peptides can also be  
C used to detect the presence of antibodies against HAV in mammalian serum.  
C The peptides can also be used to make an antibody against HAV by  
C administering the peptide to a mammal  
X  
Q Sequence 20 AA;

Query Match 95.0%; Score 19; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
b 13 VAEF 16

RESULT 22  
AB69462  
D AAB69462 standard; peptide; 21 AA.  
X  
C AAB69462;  
X  
T 20-APR-2001 (first entry)

XX Synthetic HAV P3A peptide, SEQ ID NO: 62.  
DE  
XX  
KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;  
KW antigen; major structural capsid polypeptide; HAV antibody detection.  
XX  
OS Hepatitis A virus.  
OS Synthetic.  
XX  
EN WO200105824-A2.  
XX  
PD 25-JAN-2001.  
XX  
PF 14-JUL-2000; 2000WO-US019267.  
XX  
PR 15-JUL-1999; 99US-0144412P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
FI Fields HA, Khudyakov YE;  
XX  
DR WPI; 2001-112681/12.  
XX  
PT Synthetic peptides used as antigen sources for enzyme immunoassays  
PT detecting anti-hepatitis A virus and as vaccines.  
XX  
PS Claim 22; Page 108; 130pp; English.  
XX  
CC The present sequence is one of a number of synthetic peptides which are  
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides  
CC comprise antigenic epitopes of the major structural capsid polypeptides  
CC or non-structural polypeptides of HAV with one or more glutamine  
CC molecules at the carboxy end of the peptide. The peptides are used to  
CC detect the presence of antibodies against HAV in mammalian serum, to  
CC detect the presence of HAV in a human or animal through the binding of  
CC the peptide to an antibody, to detect acute phase infection by detecting  
CC IGM antibodies in mammalian serum and detecting convalescence in a  
CC mammal. The peptides are used to detect or quantify HAV antibodies in  
CC samples in clinical or research-based assays using immunoblotting,  
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,  
CC tracking of radioactive or bioluminescent markers, chromatography or  
CC electrophoresis. The peptides are used to induce an immune response to  
CC HAV when administered to a human or animal. Glutamine at the carboxy end  
CC of the peptides enhances the IGM antibody reactivity  
XX  
SQ Sequence 21 AA;

Query Match 95.0%; Score 19; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
b 13 VAEF 16

RESULT 23  
AAB69463  
ID AAB69463 standard; peptide; 21 AA.  
XX  
AC AAB69463;  
XX  
DT 20-APR-2001 (first entry)  
XX  
DE Synthetic HAV P3A peptide, SEQ ID NO: 63.  
XX  
KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;  
KW antigen; major structural capsid polypeptide; HAV antibody detection.  
OS  
OS Hepatitis A virus.  
OS Synthetic.  
XX  
EN WO200105824-A2.



XX PD 25-JAN-2001.  
 XX PF 14-JUL-2000; 2000WO-US019267.  
 XX PR 15-JUL-1999; 99US-0144412P.  
 XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX PI Fields HA, Khudyakov YE;  
 XX DR WPI; 2001-112681/12.  
 XX PT Synthetic peptides used as antigen sources for enzyme immunoassays  
 XX PT detecting anti-hepatitis A virus and as vaccines.  
 XX PS Claim 22; Page 109; 130pp; English.  
 XX CC The present sequence is one of a number of synthetic peptides which are  
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides  
 CC comprise antigenic epitopes of the major structural capsid polypeptides  
 CC or non-structural polypeptides of HAV with one or more glutamine  
 CC molecules at the carboxy end of the peptide. The peptides are used to  
 CC detect the presence of antibodies against HAV in mammalian serum, to  
 CC detect the presence of HAV in a human or animal through the binding of  
 CC the peptide to an antibody, to detect acute phase infection by detecting  
 CC IGM antibodies in mammalian serum and detecting convalescence in a  
 CC mammal. The peptides are used to detect or quantify HAV antibodies in  
 CC samples in clinical or research-based assays using immunoblotting,  
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,  
 CC tracking of radioactive or bioluminescent markers, chromatography or  
 CC electrophoresis. The peptides are used to induce an immune response to  
 CC HAV when administered to a human or animal. Glutamine at the carboxy end  
 CC of the peptides enhances the IGM antibody reactivity  
 XX SQ Sequence 21 AA;  
 Query Match 95.0%; Score 19; DB 4; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 DB 4 VAEF 7  
 RESULT 24  
 ID AAW42946  
 AC AAW42946 standard; peptide; 25 AA.  
 DT 28-APR-1998 (first entry)  
 DE Immunogenic Hepatitis A virus peptide YK-1832.  
 XX Immunogenic peptide; immunogenic epitope; P3A protein; immune response;  
 XX antibody.  
 OS Synthetic.  
 OS Hepatitis A virus.  
 XX WO9740147-A1.  
 XX PD 30-OCT-1997.  
 XX PF 18-APR-1997; 97WO-US006891.  
 XX PR 19-APR-1996; 96US-0015644P.  
 XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX PI Fields HA, Khudyakov YE;

XX WPI; 1997-535831/49.  
 XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune  
 XX response to HAV in a mammal or to detect the presence of antibodies  
 XX against HAV in a mammal.  
 XX PS Claim 33; Page 115; 140pp; English.  
 XX CC The present immunogenic peptide corresponds to an immunogenic epitope of  
 CC the Hepatitis A virus (HAV). The peptide is substantially similar to a  
 CC portion of the amino acid sequence of the P3A protein of HAV  
 CC corresponding to amino acids 1423-1496. Compositions containing the  
 CC peptide can be used to induce an immune response to HAV in a mammal. The  
 CC peptide can also be used to detect the presence of antibodies against HAV  
 CC in mammalian serum. The peptide can also be used to make an antibody  
 CC against HAV by administering the peptide to a mammal  
 XX SQ Sequence 25 AA;  
 Query Match 95.0%; Score 19; DB 2; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 DB 3 VAEF 6  
 RESULT 25  
 ID AAB69465  
 AC AAB69465 standard; peptide; 25 AA.  
 XX AAB69465;  
 XX DT 20-APR-2001 (first entry)  
 XX DE Synthetic HAV P3A peptide, SEQ ID NO: 65.  
 XX KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;  
 XX KW antigen; major structural capsid polypeptide; HAV antibody detection.  
 XX OS Hepatitis A virus.  
 XX OS Synthetic.  
 XX PN WO200105824-A2.  
 XX PD 25-JAN-2001.  
 XX PF 14-JUL-2000; 2000WO-US019267.  
 XX PR 15-JUL-1999; 99US-0144412P.  
 XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX PI Fields HA, Khudyakov YE;  
 XX DR WPI; 2001-112681/12.  
 XX PT Synthetic peptides used as antigen sources for enzyme immunoassays  
 XX PT detecting anti-hepatitis A virus and as vaccines.  
 XX PS Claim 22; Page 110; 130pp; English.  
 XX CC The present sequence is one of a number of synthetic peptides which are  
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides  
 CC comprise antigenic epitopes of the major structural capsid polypeptides  
 CC or non-structural polypeptides of HAV with one or more glutamine  
 CC molecules at the carboxy end of the peptide. The peptides are used to  
 CC detect the presence of antibodies against HAV in mammalian serum, to  
 CC detect the presence of HAV in a human or animal through the binding of  
 CC the peptide to an antibody, to detect acute phase infection by detecting  
 CC IGM antibodies in mammalian serum and detecting convalescence in a  
 CC mammal.

XC mammal. The peptides are used to detect or quantify HAV antibodies in  
 XC samples in clinical or research-based assays using immunoblotting,  
 XC fluorescence in situ hybridisation analysis, gel-mobility shift assays,  
 XC tracking of radioactive or bioluminescent markers, chromatography or  
 XC electrophoresis. The peptides are used to induce an immune response to  
 XC HAV when administered to a human or animal. Glutamine at the carboxy end  
 XC of the peptides enhances the IGM antibody reactivity  
 XQ Sequence 25 AA;

Query Match 95.0%; Score 19; DB 4; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 QD 3 VAEF 6

## RESULT 26

AAW17832  
 ID AAW17832 standard; peptide; 28 AA.

AC AAW17832;

DT 29-JUL-1997 (first entry)

DE Human mevalonate pyrophosphate decarboxylase peptide NP71.

GW Mevalonate pyrophosphate decarboxylase; MPD; cholesterol.

DS Homo sapiens.

PN WO9714787-A1.

PD 24-APR-1997.

PF 10-OCT-1996; 96WO-EP004394.

PR 18-OCT-1995; 95US-0005652P.

PA (NOVS) NOVARTIS AG.

PI Toth MJ, Huwylar LR;

PR MPI; 1997-245104/22.

Human mevalonate pyrophosphate decarboxylase coding sequence - used for  
 screening for MPD inhibitors, which regulate and control cholesterol  
 synthesis.

Example 3; Page 10; 37pp; English.

Peptide NP71 (AAW17832) was obtd. by trypsin digestion of human liver  
 mevalonate pyrophosphate decarboxylase (MPD) (AAW17831), an enzyme of the  
 cholesterol biosynthetic pathway. The sequence of the peptide was used to  
 design PCR primers utilised in the amplification of cDNA from a rat liver  
 cDNA library. A rat MPD partial clone was obtd. and used as a probe to  
 screen a human liver cDNA library. A 1800 bp sequence (AAW66464) coding  
 for human MPD was identified

Sequence 28 AA;

Query Match 95.0%; Score 19; DB 2; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 QD 20 VAEF 23

## RESULT 27

WO200157273-A2.

AAW33910  
 ID AAW33910 standard; protein; 28 AA.  
 AC AAW33910;  
 XX 17-OCT-2001 (first entry)

DT Peptide #7947 encoded by probe for measuring placental gene expression.

DE Probe; microarray; human; placenta; antenatal diagnosis;

GW genetic disorder.

OS Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000663.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PR MPI; 2001-488897/53.

Human genome-derived single exon nucleic acid probes useful for analyzing  
 gene expression in human placenta.

Claim 27; SEQ ID NO 34179; 654pp; English.

The present invention relates to single exon nucleic acid probes (SENP;  
 see AAI31315-AA157546). The present sequence is a peptide encoded by one  
 such probe. The probes are useful for producing a microarray for  
 predicting, measuring and displaying gene expression in samples derived  
 from human placenta. The probes are useful for antenatal diagnosis of  
 human genetic disorders

Sequence 28 AA;

Query Match 95.0%; Score 19; DB 4; Length 28;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

QD 4 VAEF 7

## RESULT 28

ABG55464

ID ABG55464 standard; peptide; 28 AA.

AC ABG55464;

DT 25-FEB-2003 (first entry)

DE Human liver peptide, SEQ ID No 34112.

KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

KW hypercholesterolaemia; coronary heart disease.

OS Homo sapiens.

PN WO200157273-A2.

```

XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000664.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PF WPI; 2001-488898/53.
XX DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human adult liver.
XX PS Claim 27; SEQ ID NO 3411.2; 658pp; English.
XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX CC measuring human gene expression in a sample derived from human adult
XX CC liver, comprising one of 13109 defined nucleotide sequences given in the
XX CC specification (or complements/ fragments). The probe hybridises at high
XX CC stringency to a nucleic acid molecule expressed in the human adult liver.
XX CC (I) may be used for predicting, measuring and displaying gene expression
XX CC in samples derived from human adult liver. The genes identified may be
XX CC involved in genetic liver diseases such as cirrhosis,
XX CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX CC associated with coronary heart disease. ABG47348-ABG59930 represent human
XX CC liver single exon encoded peptides of the invention. Note: The sequence
XX CC information for this patent does not appear in the printed specification
XX CC but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX CC Sequence 28 AA;
XX SQ
XX Query Match 95.0%; Score 19; DB 4; Length 28;
XX Best Local Similarity 100.0%; Pred. No. 3.7e+02;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 2 VAEF 5
XX DB 4 VAEF 7
XX RESULT 29
XX ABG43601
XX ID ABG43601 standard; peptide; 28 AA.
XX AC ABG43601;
XX DT 19-AUG-2002 (first entry)
XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 33266.
XX KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
XX KW chronic obstructive pulmonary disease; interstitial lung disease;
XX KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX KW primary ciliary dyskinesia; pulmonary hypertension;
XX KW hyaline membrane disease.
XX OS Homo sapiens.
XX PN WO200186003-A2.

```

---

```

XX PD 15-NOV-2001.
XX PF 30-JAN-2001; 2001WO-US000665.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PF WPI; 2002-114183/15.
XX DR Spatially-addressable set of single exon nucleic acid probes, used to
XX PT measure gene expression in human lung samples.
XX PS Claim 27; SEQ ID NO 33266; 634pp; English.
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human lung comprising single exon nucleic acid probes having one of
XX CC 12614 nucleic acid sequences mentioned in the specification, or their
XX CC complements or the 12387 open reading frames derived from the 12614
XX CC probes. Also included are a microarray comprising the novel set of probes
XX CC ; the novel set of probes which hybridise at high stringency to a nucleic
XX CC acid expressed in the human lung; measuring gene expression in a sample
XX CC derived from human lung, comprising (a) contacting the array with a
XX CC collection of detectably labeled nucleic acids derived from human lung
XX CC mRNA, and (b) measuring the label detectably bound to each probe of the
XX CC array; identifying exons in a eukaryotic genome, comprising (a)
XX CC algorithmically predicting at least one exon from genomic sequences of
XX CC the eukaryote; and (b) detecting specific hybridisation of detectably
XX CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX CC having a fragment identical to the predicted exon, the probe is included
XX CC in the above mentioned microarray; assigning exons to a single gene,
XX CC comprising (a) identifying exons from genomic sequence by the method
XX CC above and (b) measuring the expression of each of the exons in several
XX CC tissues and/or cell types using hybridisation to a single exon
XX CC microarrays having a probe with the exon, where a common pattern of
XX CC expression of the exons in the tissues and/or cell types indicates that
XX CC the exons should be assigned to a single gene; a peptide comprising one
XX CC of 12011 sequences, mentioned in the specification, or encoded by the
XX CC probes/open reading frames (ORF). The probes are used for gene expression
XX CC analysis, and for identifying exons in a gene, particularly using human
XX CC lung derived mRNA and for the study of lung diseases such as asthma, lung
XX CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
XX CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
XX CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
XX CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
XX CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
XX CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
XX CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
XX CC present sequence is a peptide/protein encoded by a single exon probe of
XX CC the invention. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC -directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ
XX Sequence 28 AA;
XX Query Match 95.0%; Score 19; DB 5; Length 28;
XX Best Local Similarity 100.0%; Pred. No. 3.7e+02;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 2 VAEF 5
XX DB 4 VAEF 7

```



PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241224P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244517P.  
PR 08-NOV-2000; 2000US-024474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246509P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 08-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483426/52.

DR N-PSDB; AAK55435.  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX Claim 11; SEQ ID NO 10247; 3071pp + Sequence Listing; English.  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK67694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention  
XX Sequence 32 AA;  
SQ  
Query Match 95.0%; Score 19; DB 4; Length 32;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
Db 13 VAEF 16  
|||  
RESULT 32  
AAO09024  
ID AAO09024 standard; protein; 33 AA.  
XX AAO09024;  
AC AAO09024;  
XX 06-NOV-2001 (first entry)  
DT Human polypeptide SEQ ID NO 22916.  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.  
XX Homo sapiens.  
XX WO200164835-A2.  
XX 07-SEP-2001.  
XX 26-FEB-2001; 2001WO-US004927.  
XX 28-FEB-2000; 2000US-00515126.  
XX 18-MAY-2000; 2000US-00577409.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-514838/56.  
XX N-PSDB; AAK88955.  
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
PT and treating e.g. leukemia, inflammation and immune disorders.  
XX Claim 20; SEQ ID NO 22916; 1399pp + Sequence Listing; English.  
PS



Query Match 95.0%; Score 19; DB 2; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 Db 14 VAEF 17

RESULT 35  
 ID AAY51751 standard; protein; 35 AA.

XX AC AAY51751;  
 XX AC AAY51751;  
 XX 13-JUN-2000 (first entry)  
 XX DE H. influenzae antigenic Tbp2 peptide TBP2-21.

XX KW Transferrin receptor; Tbp1; Tbp2; immunogenic; antibacterial; vaccine;  
 XX KW diagnosis.

XX OS Haemophilus influenzae.

XX PN US6015688-A.

XX PD 18-JAN-2000.

XX PF 07-JUN-1995; 95US-00483577.

XX PR 08-NOV-1993; 93US-00148968.

XX PR 23-DEC-1993; 93US-00175116.

XX PR 08-NOV-1994; 94US-00337483.

XX PA (CONN-) CONNAUGHT LAB LTD.

XX PI Loomore S, Harkness R, Schryvers A, Gray-Owen S, Yang Y;

XX PI Murdin A, Klein M, Chong P;

XX DR WPI; 1997-052329/05.

XX PT Haemophilus truncated transferrin receptor protein analogue, Tbp2 - used  
 XX PT to induce protection against disease caused by transferrin producing  
 XX PT pathogens, or as antigen to detect Haemophilus Tfr antibodies.

XX PS Example 16; Col 39-40; 281pp; English.

XX CC This invention describes a novel isolated and purified nucleic acid (I)  
 CC encoding an immunogenic, C-terminally truncated analog of one of the  
 CC transferrin receptor proteins Tbp1 or Tbp2 of Haemophilus influenzae  
 CC which has antibacterial activity. (I) are used for recombinant production  
 CC of truncated Tbp; as probes and primers for detecting, and diagnosing  
 CC infection by, Haemophilus, also for isolating similar sequences from  
 CC other bacteria; as immunogens for vaccinating against infections caused  
 CC by bacteria that produce transferrin receptors, e.g. Haemophilus,  
 CC Neisseria or Branhamella. The truncated proteins are useful as immunogens  
 CC (as above); for diagnosing infection (as antigens in immunoassays) and  
 CC for raising antibodies, used for diagnosis of infections or for passive  
 CC immunization. AAY51695-Y51767 represent H. influenzae transferrin  
 CC receptor proteins Tbp1 and Tbp2 antigenic peptide fragments

XX SQ Sequence 35 AA;

Query Match 95.0%; Score 19; DB 2; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 Db 14 VAEF 17

RESULT 37

ID AAW54105 standard; peptide; 35 AA.

XX AC AAW54105;

XX DT 20-JUL-1998 (first entry)

XX DE Tbp2 antigenic peptide TBP2-21.

RESULT 36

AAW55788

ID AAW55788 standard; peptide; 35 AA.

XX AC AAW55788;

XX DT 14-JUL-1998 (first entry)

XX DE Human aquaporin-1 NH2-terminal amino acid sequence CHIP28.

XX KW Human; aquaporin-1; AQP-1; water channel protein; regulation;  
 XX KW osmotic change; erythrocyte; dryness; blindness; hydration; asthma;  
 XX KW secretion.

XX OS Homo sapiens.

XX FT Key

XX FT Location/Qualifiers

XX FT Misc-difference 30

XX FT /label= Unknown

XX FT /note= "not specified but is given as Ser in the full  
 length protein given in AAW55786"

XX PN US5741671-A.

XX PD 21-APR-1998.

XX PF 06-JUN-1995; 95US-00468763.

XX PR 17-AUG-1992; 92US-00930168.

XX PR 24-FEB-1995; 95US-00393996.

XX PA (UYJO ) UNIV JOHNS HOPKINS.

XX PI Agre PC;

XX DR WPI; 1998-260501/23.

XX PT Polynucleotide(s) encoding water channel protein Aquaporin-1 - are useful  
 XX PT for recombinant production of protein for activity studies.

XX PS Example 2; Col 20; 48pp; English.

XX CC The present sequence represents the NH2-terminal amino acid sequence of  
 CC aquaporin-1 (AQP-1), isolated from human erythrocytes. AQP-1 is a water  
 CC channel protein. Water channel proteins regulate the passage of water in  
 CC and out of cells, in response to osmotic changes. The DNA encoding AQP-1  
 CC is useful for the recombinant production of AQP-1, found in mammalian  
 CC erythrocytes, and is useful in the study to identify reagents which  
 CC enhance or inhibit water channel function. This can lead to therapeutics  
 CC which enhance secretion e.g. in the case of dryness of eyes which can  
 CC lead to blindness or to hydrate large respiratory airways, as their  
 CC dryness can precipitate asthma

XX SQ Sequence 35 AA;

Query Match 95.0%; Score 19; DB 2; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 Db 14 VAEF 17





CC fragment) from strains of Haemophilus influenzae. This sequence  
 CC corresponds to an epitope from the H. influenzae transferrin receptor  
 CC protein Tbp2. The antibodies may be used for preventing and treating  
 CC infections and disorders caused by H. influenzae, including bacterial  
 CC meningitis, otitis media, epiglottitis, pneumonia and tracheobronchitis.  
 CC The antibodies may also be used to detect the presence of H. influenzae  
 CC proteins in samples according to standard methodologies (e.g. enzyme  
 CC linked immunosorbent assay (ELISA)) and hence diagnose infections  
 XX  
 SQ Sequence 35 AA;

Query Match 95.0%; Score 19; DB 3; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+02; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0;

QY 2 VAEF 5  
 Db 14 VAEF 17  
 |||||

RESULT 40  
 AAW27782  
 ID AAW27782 standard; protein; 36 AA.  
 XX AC AAW27782;  
 XX DT 21-JUL-1998 (first entry)  
 XX DE UDP-N-acetylglucosamine 1-carboxyvinyltransferase.  
 XX KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;  
 XX KW Staphylococcal gene; regulatory element; bacterial gene expression;  
 XX KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;  
 XX KW toxic shock syndrome.  
 XX OS Staphylococcus aureus.  
 XX PH Key Location/Qualifiers  
 XX FT Misc-difference 34 /note= "not specified"  
 XX FT  
 XX PN WO9730070-A1.  
 XX PD 21-AUG-1997.  
 XX PF 19-FEB-1997; 97WO-US002318.  
 XX PR 20-FEB-1996; 96US-0011888P.  
 XX PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 XX PI Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;  
 XX PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;  
 XX PI WPI; 1997-424969/39.  
 XX DR N-PSDB; AAT83751.  
 XX FT Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used to  
 XX FT isolate antimicrobial compounds, and in vaccines against S. aureus  
 XX FT infection.  
 XX PT  
 XX PS Claim 6; Page 272; 989pp; English.  
 XX CC The present sequence represents a Staphylococcus aureus protein, that,  
 XX CC based on homology with a Bacillus subtilis protein, is believed to be a  
 XX CC probable UDP-N-acetylglucosamine 1-carboxyvinyltransferase (enoylpyruvate  
 XX CC transferase, UDP-N-acetylglucosamine enoylpyruvyl transferase). The DNA  
 XX CC sequence was isolated from a library of clones of S. aureus WCUH 29 in  
 XX CC Escherichia coli. The DNA sequence can be used in the construction of  
 XX CC ribozymes and antisense sequences to control the expression of  
 XX CC Staphylococcal genes. The DNA sequence is also useful as a source of  
 XX CC regulatory elements for the control of bacterial gene expression. The  
 XX CC present protein may be used to produce vaccines to enable a host to

CC produce specific antibodies with antibacterial action. These vaccines and  
 CC antibodies would protect a host against invasion by S. aureus, and  
 CC conditions relating to staphylococcal infection, e.g. Staphylococcal food  
 CC poisoning, scaled skin syndrome, and toxic shock syndrome  
 XX  
 SQ Sequence 36 AA;

Query Match 95.0%; Score 19; DB 2; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 4.8e+02; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0;

QY 2 VAEF 5  
 Db 3 VAEF 6  
 |||||

RESULT 41  
 AAM99833  
 ID AAM99833 standard; protein; 37 AA.  
 XX AC AAM99833;  
 XX DT 07-JAN-2002 (first entry)  
 XX DE Human excretory related polypeptide SEQ ID NO 570.  
 XX KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
 XX KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
 XX KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;  
 XX KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
 XX KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
 XX KW antiparasitic; cardiast; immune disorder; cardiovascular disorder;  
 XX KW neurological disease; infection; nephrotropic; gene therapy; vaccine;  
 XX KW excretory system.  
 XX OS Homo sapiens.  
 XX PN WO200155313-A2.  
 XX PD 02-AUG-2001.  
 XX PF 17-JAN-2001; 2001WO-US001323.  
 XX PR 31-JAN-2000; 2000US-0179065P.  
 XX PR 04-FEB-2000; 2000US-0180628P.  
 XX PR 24-FEB-2000; 2000US-0184664P.  
 XX PR 02-MAR-2000; 2000US-0186350P.  
 XX PR 16-MAR-2000; 2000US-0189874P.  
 XX PR 17-MAR-2000; 2000US-0190076P.  
 XX PR 18-APR-2000; 2000US-0198123P.  
 XX PR 19-MAY-2000; 2000US-0205515P.  
 XX PR 07-JUN-2000; 2000US-0209467P.  
 XX PR 28-JUN-2000; 2000US-0214886P.  
 XX PR 30-JUN-2000; 2000US-0215135P.  
 XX PR 07-JUL-2000; 2000US-0216647P.  
 XX PR 07-JUL-2000; 2000US-0216880P.  
 XX PR 11-JUL-2000; 2000US-0217487P.  
 XX PR 11-JUL-2000; 2000US-0217496P.  
 XX PR 14-JUL-2000; 2000US-0218290P.  
 XX PR 26-JUL-2000; 2000US-0220963P.  
 XX PR 26-JUL-2000; 2000US-0220964P.  
 XX PR 14-AUG-2000; 2000US-0224518P.  
 XX PR 14-AUG-2000; 2000US-0224519P.  
 XX PR 14-AUG-2000; 2000US-0225213P.  
 XX PR 14-AUG-2000; 2000US-0225214P.  
 XX PR 14-AUG-2000; 2000US-0225266P.  
 XX PR 14-AUG-2000; 2000US-0225267P.  
 XX PR 14-AUG-2000; 2000US-0225268P.  
 XX PR 14-AUG-2000; 2000US-0225270P.  
 XX PR 14-AUG-2000; 2000US-0225477P.  
 XX PR 14-AUG-2000; 2000US-0225757P.  
 XX PR 14-AUG-2000; 2000US-0225758P.  
 XX PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226888P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232377P.  
PR 14-SEP-2000; 2000US-0232388P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241896P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246533P.  
PR 08-NOV-2000; 2000US-0246534P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251388P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-465569/50.  
DR N-PSDE; AA198806.  
XX  
PT Isolated nucleic acid molecule encoding excretory system antigen is used  
XX in preventing, treating or ameliorating a medical condition.  
PS Claim 11; SEQ ID NO 570; 574pp + Sequence Listing; English.  
XX  
CC The invention relates to novel excretory system related human  
CC polynucleotides (AA198567-AA199503) and the encoded proteins (AAW99594-  
CC AA99913) useful for preventing, treating or ameliorating medical  
CC conditions e.g. by protein or gene therapy, especially disorders related  
CC to the excretory system. The genes are isolated from a range of human  
CC tissues disclosed in the specification. The nucleic acids, proteins,  
CC antibodies and (ant)agonists are useful in the diagnosis, treatment and  
CC prevention of: (a) cancer, e.g. breast and ovarian cancer and other  
CC cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal  
CC tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's  
CC disease, allergies, autoimmune haemolytic anaemia, autoimmune  
CC thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,  
CC rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders  
CC such as myocardial ischaemias; (d) wound healing; (e) neurological  
CC diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases  
CC such as viral, bacterial, fungal and parasitic infections. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
CC Sequence 37 AA;

Query Match 95.0%; Score 19; DB 4; Length 37;  
Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

Db 14 VAEP 17  
RESULT 42  
AA042648 standard; protein; 37 AA.  
AC AA042648;  
XX 22-OCT-2001 (first entry)  
XX Human kidney related polypeptide SEQ ID NO 517.  
XX Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
KW neurotropic; neuroprotective; antibacterial; virucide; fungicide;  
KW ophthalmological; antiallergic; hepatotropic; antidiabetic;  
KW antiinflammatory; antitumor; cancer; immune disorder; cardiovascular disorder;  
KW gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection.  
XX Homo sapiens.  
OS WO200155323-A2.  
XX 02-AUG-2001.  
XX 17-JAN-2001; 2001WO-US001343.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205151P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225472P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0234984P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236227P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239835P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.

17-NOV-2000; 2000US-0249244P.  
 17-NOV-2000; 2000US-0249245P.  
 17-NOV-2000; 2000US-0249246P.  
 17-NOV-2000; 2000US-0249265P.  
 17-NOV-2000; 2000US-0249297P.  
 17-NOV-2000; 2000US-0249299P.  
 17-NOV-2000; 2000US-0249300P.  
 01-DEC-2000; 2000US-0250160P.  
 01-DEC-2000; 2000US-0250391P.  
 05-DEC-2000; 2000US-0251030P.  
 05-DEC-2000; 2000US-0251988P.  
 05-DEC-2000; 2000US-0256719P.  
 06-DEC-2000; 2000US-0251479P.  
 08-DEC-2000; 2000US-0251856P.  
 08-DEC-2000; 2000US-0251868P.  
 08-DEC-2000; 2000US-0251869P.  
 08-DEC-2000; 2000US-0251989P.  
 08-DEC-2000; 2000US-0251990P.  
 11-DEC-2000; 2000US-0254097P.  
 05-JAN-2001; 2001US-0259678P.  
 (HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;  
 WPI; 2001-488784/53.  
 N-PSDB; AAI63202.

New isolated nucleic acids and polypeptides, useful for diagnosing,  
 treating and/or preventing human diseases and disorders.

Claim 11; SEQ ID NO 517; 564pp + Sequence Listing; English.

The invention relates to novel kidney related polynucleotides (AAI62971-  
 AAI63793) and the encoded polypeptides (AAI42417-AAI42691) collectively  
 known as kidney antigens and the use of such kidney antigens for  
 detecting disorders of the kidney, especially kidney cancer and kidney  
 cancer metastases. The polynucleotides and proteins are also useful for  
 preventing, treating or ameliorating medical conditions e.g. by protein  
 or gene therapy. The genes are isolated from a range of human tissues  
 disclosed in the specification. The nucleic acids, proteins, antibodies  
 and (ant)agonists are useful in the diagnosis, treatment and prevention  
 of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the  
 adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
 lung, or urogenital; (b) immune disorders e.g. Addison's disease,  
 allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
 diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 arthritis and ulcerative colitis; (c) cardiovascular disorders such as  
 myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.  
 cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,  
 bacterial, fungal and parasitic infections. Note: The sequence data for  
 this patent did not form part of the printed specification, but was  
 obtained in electronic format directly from WIPO at  
 ftp.wipo.int/pub/published\_pct\_sequences

Sequence 37 AA;

Query Match 95.0%; Score 19; DB 4; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

Db 14 VAEF 17

RESULT 43

ID AAI17544  
 AA17544 standard; protein; 43 AA.

AC AA17544;

DT 12-OCT-2001 (first entry)

XX Peptide #3978 encoded by probe for measuring cervical gene expression.  
 DE Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 OS Homo sapiens.  
 PN WO200157278-A2.  
 PD 09-AUG-2001.  
 XX 30-JAN-2001; 2001WO-US000670.  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 WPI; 2001-488901/53.

Human genome-derived single exon nucleic acid probes useful for analyzing  
 gene expression in human cervical epithelial cells.

Claim 27; SEQ ID NO 22370; 487pp; English.

The present invention relates to human single exon nucleic acid probes  
 (SENPs) see AAI0068-AAI28459). The present sequence is a peptide encoded  
 by one such probe. The SENPs are derived from human Hela cells. The SENPs  
 can be used to produce a single exon microarray, which can be used for  
 measuring human gene expression in a sample derived from human cervical  
 epithelial cells. By measuring gene expression, the probes are therefore  
 useful in grading and/or staging of diseases of the cervix, notably  
 cervical cancer. Note: The sequence data for this patent did not form  
 part of the printed specification, but was obtained in electronic format  
 directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 43 AA;

Query Match 95.0%; Score 19; DB 4; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

Db 40 VAEF 43

RESULT 44

ID ABB31362  
 ABB31362 standard; peptide; 43 AA.

AC ABB31362;

DT 01-FEB-2002 (first entry)

DE Peptide #4013 encoded by breast cell single exon nucleic acid probe.

KW Human; microarray; single exon probe; gene expression; breast; disease;  
 cancer.

OS Homo sapiens.

PN WO200157271-A2.

XX 09-AUG-2001.

```
XX PF 30-JAN-2001; 2001WO-US000662.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLB-) MOLECULAR DYNAMICS INC.
XX XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX FI MPI; 2001-496933/54.
XX DR
XX XX New spatially-addressable set of single exon nucleic acid probes, useful
XX PT for measuring gene expression in sample derived from human breast,
XX PT comprises number of single exon nucleic acid probes.
XX PS Claim 27; SEQ ID NO 14330; 327pp + Sequence Listing; English.
XX XX
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and BT 474 cells. The method involves contacting the
XX CC probes with a collection of detectably labelled nucleic acids derived
XX CC from mRNA of human breast, and then measuring the label bound to each
XX CC probe of the microarray. The probes are useful for verifying the
XX CC expression of regions of genomic DNA predicted to encode proteins. They
XX CC are useful for gene discovery, and for determining predisposition and/or
XX CC prognosing breast disease. Gene expression analysis is useful for
XX CC assessing the toxicity of chemical agents on cells. The microarray of
XX CC this invention presents a far greater diversity of probes for measuring
XX CC gene expression, with far less bias than expressed sequence tag
XX CC microarrays. The method is suitable for rapid production of functional
XX CC information from genomic sequence. The present sequence is a peptide
XX CC encoded by a single exon nucleic acid probe of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 43 AA;
XX
XX Query Match 95.0%; Score 19; DB 4; Length 43;
XX Best Local Similarity 100.0%; Pred. No. 5.9e+02;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 VAEF 5
XX DB 40 VAEF 43
XX
XX RESULT 45
XX ABB21904
XX ID ABB21904 standard; protein; 43 AA.
XX AC
XX AC ABB21904;
XX XX
XX DT 23-JAN-2002 (first entry)
XX DE
XX DE Protein #3903 encoded by probe for measuring heart cell gene expression.
XX KW Human; gene expression; heart; microarray; vascular system;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease.
XX OS Homo sapiens.
XX XX
XX XX WO200157274-A2.
XX PN
XX XX 09-AUG-2001.
XX PD
XX XX
XX Query Match 95.0%; Score 19; DB 4; Length 43;
XX Best Local Similarity 100.0%; Pred. No. 5.9e+02;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 VAEF 5
XX DB 40 VAEF 43
XX
XX RESULT 46
XX ABB51414
XX ID ABB51414 standard; peptide; 43 AA.
XX AC
XX AC ABB51414;
XX XX
XX DT 25-FEB-2003 (first entry)
XX DE
XX DE Human liver peptide, SEQ ID No 30062.
XX KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
XX KW hypercholesterolaemia; coronary heart disease.
XX OS Homo sapiens.
XX XX
XX XX WO200157273-A2.
XX PN
XX XX 09-AUG-2001.
XX PD
XX XX
XX 30-JAN-2001; 2001WO-US000664.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLB-) MOLECULAR DYNAMICS INC.
XX XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX FI MPI; 2001-498899/53.
XX DR
XX XX Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts.
XX PT
XX XX Claim 15; SEQ ID NO 23674; 530pp; English.
XX PS
XX XX The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart (see
XX CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
XX CC probe. The probes may be used for predicting, measuring and displaying
XX CC gene expression in samples derived from the human heart via microarrays.
XX CC By measuring gene expression, the probes are useful for predicting,
XX CC diagnosing, grading, staging, monitoring and prognosing diseases of the
XX CC human heart and vascular system e.g. cardiovascular disease,
XX CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 43 AA;
XX
XX Query Match 95.0%; Score 19; DB 4; Length 43;
XX Best Local Similarity 100.0%; Pred. No. 5.9e+02;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 VAEF 5
XX DB 40 VAEF 43
XX
```

(MOLE-) MOLECULAR DYNAMICS INC.  
 Penn SG, Hanzel DK, Chen W, Rank DR;  
 WPI; 2001-48898/53.  
 Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human adult liver.  
 Claim 27; SEQ ID NO 30062; 658pp; English.  
 The invention relates to a single exon nucleic acid probe (SENP) (I) for measuring human gene expression in a sample derived from human adult liver, comprising one of 13109 defined nucleotide sequences given in the specification (or complements/ fragments). The probe hybridizes at high stringency to a nucleic acid molecule expressed in the human adult liver. (I) may be used for predicting, measuring and displaying gene expression in samples derived from human adult liver. The genes identified may be involved in genetic liver diseases such as cirrhosis, hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is associated with coronary heart disease. ABG47348-ABG59930 represent human liver single exon encoded peptides of the invention. Note: The sequence information for this patent does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 Sequence 43 AA;  
 Query Match 95.0%; Score 19; DB 4; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 ||||  
 40 VAEF 43  
 RESULT 47  
 LAM05214  
 ID AAM05214 standard; protein; 43 AA.  
 LAM05214;  
 09-OCT-2001 (first entry)  
 Peptide #3896 encoded by probe for measuring breast gene expression.  
 Probe; human; breast disease; breast cancer; development disorder; inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
 Homo sapiens.  
 WO200157270-A2.  
 09-AUG-2001.  
 29-JAN-2001; 2001WO-US000661.  
 04-FEB-2000; 2000US-0180312P.  
 26-MAY-2000; 2000US-0207456P.  
 30-JUN-2000; 2000US-00608408.  
 03-AUG-2000; 2000US-00632366.  
 21-SEP-2000; 2000US-0234687P.  
 27-SEP-2000; 2000US-0236359P.  
 04-OCT-2000; 2000GB-00024263.  
 (MOLE-) MOLECULAR DYNAMICS INC.  
 Penn SG, Hanzel DK, Chen W, Rank DR;  
 WPI; 2001-476286/51.

Novel single exon nucleic acid probe used to measuring gene expression in a human breast.  
 Claim 27; SEQ ID NO 13954; 322pp; English.  
 The present invention relates to novel single exon nucleic acid probes (see AAI00010-AA110067). The present sequence is a peptide encoded by one such probe. The probes are useful for measuring human gene expression in a human breast sample, where the probe hybridizes at high stringency to a nucleic acid expressed in the human breast. The probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human breast, particularly those diseases with polygenic aetiology. The diseases include: breast cancer, disorders of development, inflammatory diseases of the breast, fibrocystic changes, proliferative breast disease and non-carcinoma tumours. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 Sequence 43 AA;  
 Query Match 95.0%; Score 19; DB 4; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 ||||  
 40 VAEF 43  
 RESULT 48  
 AAB27178  
 ID AAB27178 standard; protein; 47 AA.  
 AAB27178;  
 27-FEB-2001 (first entry)  
 RSV partial protein sequence SEQ ID NO: 26.  
 Negative stranded RNA virus; vaccine; attenuated virus; RSV; PIV; measles; respiratory syncytial virus; parainfluenza virus.  
 Respiratory syncytial virus.  
 WO2000061737-A2.  
 19-OCT-2000.  
 12-APR-2000; 2000WO-US009695.  
 13-APR-1999; 99US-0129006P.  
 (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 Murphy BR, Collins PL, Durbin AP, Skiadopoulos MH;  
 WPI; 2000-687044/67.  
 Producing attenuated negative stranded RNA virus vaccines from cloned sequences, useful for immunizing against e.g. respiratory syncytial virus, human parainfluenza virus, Sendai virus Newcastle disease virus, mumps virus and measles virus.  
 Example 1; Page 62; 137pp; English.  
 The present invention is concerned with producing vaccines against negative stranded RNA viruses. These viruses include measles, respiratory syncytial virus (RSV) and parainfluenza virus (PIV) in particular. The method of the invention comprises the production of a mutated form of the virus which attenuates the strain and enables it to be used as a vaccine. The present sequence comprises a partial viral protein sequence

```

SQ Sequence 47 AA;
Query Match          95.0%; Score 19; DB 3; Length 47;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
      ||||
Db      34 VAEF 37

RESULT 49
AAM18541
ID AAM18541 standard; protein; 48 AA.
XX
AC AAM18541;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #4975 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
FN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488901/53.
XX
Human genome-derived single exon nucleic acid probes useful for analyzing
gene expression in human cervical epithelial cells.
XX
Claim 27; SEQ ID NO 23367; 487pp; English.
XX
The present invention relates to human single exon nucleic acid probes
(SENP; see AAI10068-AAI28459). The present sequence is a peptide encoded
by one such probe. The SENPs are derived from human HeLa cells. The SENPs
can be used to produce a single exon microarray, which can be used for
measuring human gene expression in a sample derived from human cervical
epithelial cells. By measuring gene expression, the probes are therefore
useful in grading and/or staging of diseases of the cervix, notably
cervical cancer. Note: The sequence data for this patent did not form
part of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 48 AA;
Query Match          95.0%; Score 19; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
      ||||
Db      34 VAEF 37

RESULT 50
ABB37578
ID ABB37578 standard; peptide; 48 AA.
XX
AC ABB37578;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #5084 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
FN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-483447/52.
XX
Human genome-derived single exon nucleic acid probes useful for analyzing
gene expression in human fetal liver.
XX
Claim 27; SEQ ID NO 30213; 639pp + Sequence Listing; English.
XX
The invention relates to a single exon nucleic acid probe for measuring
human gene expression in a sample derived from human foetal liver. The
single exon nucleic acid probes may be used for predicting, measuring and
displaying gene expression in samples derived from human fetal liver. The
present sequence is a peptide encoded by a single exon nucleic acid probe
of the invention. Note: The sequence data for this patent did not form
part of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 48 AA;
Query Match          95.0%; Score 19; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
      ||||
Db      3 VAEF 6

RESULT 51
AAM30996
ID AAM30996 standard; protein; 48 AA.
XX
AC AAM30996;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #5033 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
```

X WO200157272-A2.  
 N  
 D 09-AUG-2001.  
 X  
 F 30-JAN-2001; 2001WO-US000663.  
 X  
 R 04-FEB-2000; 2000US-0180312P.  
 R 26-MAY-2000; 2000US-0207456P.  
 R 30-JUN-2000; 2000US-00608408.  
 R 03-AUG-2000; 2000US-00632366.  
 R 21-SEP-2000; 2000US-0234687P.  
 R 27-SEP-2000; 2000US-0236359P.  
 R 04-OCT-2000; 2000GB-00024263.  
 X (MOLE-) MOLECULAR DYNAMICS INC.  
 A  
 X Penn SG, Hanzel DK, Chen W, Rank DR;  
 I WPI; 2001-488997/53.  
 X  
 R Human genome-derived single exon nucleic acid probes useful for analyzing  
 T gene expression in human placenta.  
 T  
 X Claim 27; SEQ ID NO 31265; 654pp; English.  
 S  
 X The present invention relates to single exon nucleic acid probes (SENP;  
 C see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
 C such probe. The probes are useful for producing a microarray for  
 C predicting, measuring and displaying gene expression in samples derived  
 C from human placenta. The probes are useful for antenatal diagnosis of  
 C human genetic disorders  
 X  
 X Sequence 48 AA;  
 Q  
 Query Match 95.0%; Score 19; DB 4; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Y 2 VAEF 5  
 b ||||  
 3 VAEF 6  
 RESULT 52  
 BB22868  
 D ABB22868 standard; protein; 48 AA.  
 X  
 C ABB22868;  
 X  
 T 23-JAN-2002 (first entry)  
 X  
 E Protein #4867 encoded by probe for measuring heart cell gene expression.  
 X  
 W Human; gene expression; heart; microarray; vascular system;  
 W cardiovascular disease; hypertension; cardiac arrhythmia;  
 W congenital heart disease.  
 X  
 X Homo sapiens.  
 S  
 X WO200157274-A2.  
 N  
 N 09-AUG-2001.  
 D  
 X 30-JAN-2001; 2001WO-US000666.  
 F  
 R 04-FEB-2000; 2000US-0180312P.  
 R 26-MAY-2000; 2000US-0207456P.  
 R 30-JUN-2000; 2000US-00608408.  
 R 03-AUG-2000; 2000US-00632366.  
 R 21-SEP-2000; 2000US-0234687P.  
 R 27-SEP-2000; 2000US-0236359P.  
 R 04-OCT-2000; 2000GB-00024263.  
 X (MOLE-) MOLECULAR DYNAMICS INC.  
 A  
 X Penn SG, Hanzel DK, Chen W, Rank DR;  
 I WPI; 2001-488997/53.  
 X  
 R Human genome-derived single exon nucleic acid probes useful for analyzing  
 T gene expression in human bone marrow.  
 T  
 X

XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488997/53.  
 DR  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT hearts.  
 XX  
 PS Claim 15; SEQ ID NO 24638; 530pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 48 AA;  
 Query Match 95.0%; Score 19; DB 4; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 Db ||||  
 3 VAEF 6  
 RESULT 53  
 AAAM70687  
 ID AAAM70687 standard; protein; 48 AA.  
 XX  
 AC AAAM70687;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 30993.  
 XX  
 KW Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157276-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488997/53.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human bone marrow.



XX Example 4; SEQ ID NO 30393; 658pp + Sequence Listing; English.

PS The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is a

CC protein encoded by one of the probes of the invention

XX

XX Sequence 48 AA;

SQ

Query Match 95.0%; Score 19; DB 4; Length 48;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
| | | |  
Db 3 VAEF 6

RESULT 54

ABG52389

ID ABG52389 standard; peptide; 48 AA.

XX AC ABG52389;

XX

XX 25-FEB-2003 (first entry)

DT Human liver peptide, SEQ ID NO 31037.

XX

DE Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

XX hypercholesterolaemia; coronary heart disease.

KW

XX Homo sapiens.

OS

XX WO200157273-A2.

PN

XX 09-AUG-2001.

PD

XX 30-JAN-2001; 2001WO-US000664.

PF

XX 04-FEB-2000; 2000US-0180312P.

PR

XX 26-MAY-2000; 2000US-0207456P.

PR

XX 30-JUN-2000; 2000US-00608408.

PR

XX 03-AUG-2000; 2000US-00632366.

PR

XX 21-SEP-2000; 2000US-0234687P.

PR

XX 27-SEP-2000; 2000US-0236359P.

PR

XX 04-OCT-2000; 2000GB-00024263.

XX

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI

XX WPI; 2001-488898/53.

DR

XX Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human adult liver.

PT

XX Claim 27; SEQ ID NO 31037; 658pp; English.

PS

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for

CC measuring human gene expression in a sample derived from human adult

CC liver, comprising one of 13109 defined nucleotide sequences given in the

CC specification (or complements/ fragments). The probe hybridises at high

CC stringency to a nucleic acid molecule expressed in the human adult liver.

CC (I) may be used for predicting, measuring and displaying gene expression

CC in samples derived from human adult liver. The genes identified may be

CC involved in genetic liver diseases such as cirrhosis,

CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is

CC associated with coronary heart disease. ABG47348-ABG59930 represent human

CC liver single exon encoded peptides of the invention. Note: The sequence

CC information for this patent does not appear in the printed specification

CC but was obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX

XX Sequence 48 AA;

SQ

Query Match 95.0%; Score 19; DB 4; Length 48;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
| | | |  
Db 3 VAEF 6

RESULT 55

AAM06107

ID AAM06107 standard; protein; 48 AA.

XX AC AAM06107;

XX

XX 09-OCT-2001 (first entry)

DT

XX Peptide #4789 encoded by probe for measuring breast gene expression.

DE

XX Probe; human; breast disease; breast cancer; development disorder;

XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.

KW

XX Homo sapiens.

OS

XX WO200157270-A2.

PN

XX 09-AUG-2001.

PD

XX 29-JAN-2001; 2001WO-US000661.

PF

XX 04-FEB-2000; 2000US-0180312P.

PR

XX 26-MAY-2000; 2000US-0207456P.

PR

XX 30-JUN-2000; 2000US-00608408.

PR

XX 03-AUG-2000; 2000US-00632366.

PR

XX 21-SEP-2000; 2000US-0234687P.

PR

XX 27-SEP-2000; 2000US-0236359P.

PR

XX 04-OCT-2000; 2000GB-00024263.

XX

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI

XX WPI; 2001-476286/51.

DR

XX Novel single exon nucleic acid probe used to measuring gene expression in

PT a human breast.

PT

XX Claim 27; SEQ ID NO 14947; 322pp; English.

PS

XX The present invention relates to novel single exon nucleic acid probes

CC (see AA100010-AA110067). The present sequence is a peptide encoded by one

CC such probe. The probes are useful for measuring human gene expression in

CC a human breast sample, where the probe hybridises at high stringency to a

CC nucleic acid expressed in the human breast. The probes are useful for

CC predicting, diagnosing, grading, staging, monitoring and prognosing

CC diseases of the human breast, particularly those diseases with polygenic

CC aetiology. The diseases include: breast cancer, disorders of development,

CC inflammatory diseases of the breast, fibrocystic changes, proliferative

CC breast disease and non-carcinoma tumours. Note: The sequence data for

CC this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX

XX Sequence 48 AA;

SQ

Query Match 95.0%; Score 19; DB 4; Length 48;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Y      2 VAEF 5
b      3 VAEF 6

RESULT 56
AO10565
D AAO10565 standard; protein; 53 AA.
X C AAO10565;
X X
X T 06-NOV-2001 (first entry)
X E Human polypeptide SEQ ID NO 24457.
X W Human; cytokine; cell proliferation; cell differentiation; gene therapy;
W vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
W tissue growth factor; immunomodulatory; cancer; leukaemia;
W nervous system disorders; arthritis; inflammation.
X X
X S Homo sapiens.
X N WO200164835-A2.
X D 07-SEP-2001.
X F 26-FEB-2001; 2001WO-US004927.
X R 28-FEB-2000; 2000US-00515126.
X R 18-MAY-2000; 2000US-00577409.
X A (HYSE-) HYSEQ INC.
X X Tang YT, Liu C, Drmanac RT;
X T WPI; 2001-514838/56.
X R N-PSDB; AAI90496.
X S Isolated nucleic acids and polypeptides, useful for preventing diagnosing
X and treating e.g. leukemia, inflammation and immune disorders.
X S Claim 20; SEQ ID NO 24457; 1399pp + Sequence Listing; English.
X X The invention relates to human polynucleotides (AAI79941-AAI93841) and
X the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
X cytokine, cell proliferation or cell differentiation or which may induce
X production of other cytokines in other cell populations. The
X polynucleotides and polypeptides are useful in gene therapy, vaccines or
X peptide therapy. The polypeptides have various cytokine-like activities,
X e.g. stem cell growth factor activity, haematopoiesis regulating
X activity, tissue growth factor activity, immunomodulatory activity and
X activin/inhibin activity and may be useful in the diagnosis and/or
X treatment of cancer, leukaemia, nervous system disorders, arthritis and
X inflammation. Note: The sequence data for this patent did not form part
X of the printed specification, but was obtained in electronic format
X directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
X S Sequence 53 AA;
X Query Match 95.0%; Score 19; DB 4; Length 53;
X Best Local Similarity 100.0%; Pred. No. 7.3e+02;
X Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
X Y 2 VAEF 5
X 33 VAEF 36

RESULT 57
AO05247
D AAO05247 standard; protein; 53 AA.
X X
X T 06-NOV-2001 (first entry)
X E Human polypeptide SEQ ID NO 19139.
X W Human; cytokine; cell proliferation; cell differentiation; gene therapy;
W vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
W tissue growth factor; immunomodulatory; cancer; leukaemia;
W nervous system disorders; arthritis; inflammation.
X X
X S Homo sapiens.
X N WO200164835-A2.
X D 07-SEP-2001.
X F 26-FEB-2001; 2001WO-US004927.
X R 28-FEB-2000; 2000US-00515126.
X R 18-MAY-2000; 2000US-00577409.
X A (HYSE-) HYSEQ INC.
X X Tang YT, Liu C, Drmanac RT;
X T WPI; 2001-514838/56.
X R N-PSDB; AAI85178.
X S Isolated nucleic acids and polypeptides, useful for preventing diagnosing
X and treating e.g. leukemia, inflammation and immune disorders.
X S Claim 20; SEQ ID NO 19139; 1399pp + Sequence Listing; English.
X X The invention relates to human polynucleotides (AAI79941-AAI93841) and
X the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
X cytokine, cell proliferation or cell differentiation or which may induce
X production of other cytokines in other cell populations. The
X polynucleotides and polypeptides are useful in gene therapy, vaccines or
X peptide therapy. The polypeptides have various cytokine-like activities,
X e.g. stem cell growth factor activity, haematopoiesis regulating
X activity, tissue growth factor activity, immunomodulatory activity and
X activin/inhibin activity and may be useful in the diagnosis and/or
X treatment of cancer, leukaemia, nervous system disorders, arthritis and
X inflammation. Note: The sequence data for this patent did not form part
X of the printed specification, but was obtained in electronic format
X directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
X S Sequence 53 AA;
X Query Match 95.0%; Score 19; DB 4; Length 53;
X Best Local Similarity 100.0%; Pred. No. 7.3e+02;
X Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
X Y 2 VAEF 5
X 33 VAEF 36

RESULT 58
AAM87390
D AAM87390 standard; protein; 56 AA.
X X
X T 07-NOV-2001 (first entry)
X E Human immune/haematopoietic antigen SEQ ID NO:14983.
X W Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
W cytostatic; gene therapy; vaccine; metastasis.
X X
X S Homo sapiens.

```

XX WO200157182-A2.  
PN 09-AUG-2001.  
XX 17-JAN-2001; 2001WC-US001354.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214866P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241836P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246534P.  
PR 08-NOV-2000; 2000US-0246535P.  
PR 08-NOV-2000; 2000US-0246536P.  
PR 08-NOV-2000; 2000US-0246537P.  
PR 08-NOV-2000; 2000US-0246538P.  
PR 08-NOV-2000; 2000US-0246539P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249255P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX

I Rosen CA, Barash SC, Ruben SM;  
 X WPI; 2001-483426/52.  
 R N-PSDB; AAK60171.  
 T Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 X useful for preventing, diagnosing and/or treating cancers and metastasis.  
 T Claim 11; SEQ ID NO 14993; 3071pp + Sequence Listing; English.  
 S  
 X AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 C amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic  
 C activity, and can be used in gene therapy and vaccine production. (I)  
 C proteins and polynucleotides may be used in the prevention, diagnosis and  
 C treatment of diseases associated with inappropriate (I) expression. For  
 C example, they may be used to treat disorders associated with decreased  
 C expression by rectifying mutations or deletions in a patient's genome  
 C that affect the activity of (I) by expressing inactive proteins or to  
 C supplement the patient's own production of (I). Additionally, (I)  
 C polynucleotides may be used to produce the secreted (I), by inserting the  
 C nucleic acids into a host cell and culturing the cell to express the  
 C protein. (I) proteins and polynucleotides may be used to prevent,  
 C diagnose and treat immune/hematopoietic-related diseases, especially  
 C cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
 C to AAK87694 represent human immune/hematopoietic antigen genomic  
 C sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
 C represent sequences used in the exemplification of the present invention  
 X Sequence 56 AA;  
 X

Query Match 95.0%; Score 19; DB 4; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
 b 1 VAEF 4

RESULT 59

BP97126  
 D ABP97126 standard; peptide; 56 AA.  
 C ABP97126;  
 X  
 X 24-JUN-2003 (first entry)  
 X Human matrix metalloproteinase 7 cleavage region peptide SEQ ID NO:4.  
 E  
 X Human; matrix metalloproteinase; MMP; anticancer; wound healing;  
 W matrix metalloproteinase inhibitor; antitumor; antiangiogenic; cardiant;  
 W vascular endothelial growth factor inhibitor; VEGF inhibitor; cytostatic;  
 W vulnary; cerebroprotective; antidiabetic; ophthalmological; tumour;  
 W dermatological; metastatic; non-metastatic; vascularised; heart disease;  
 W non-vascularised; surgical incision; chronic wound; stroke; angiogenesis;  
 W macular degeneration; diabetic retinopathy; cleavage region.  
 S Homo sapiens.  
 X  
 X WO2003018748-A2.  
 X  
 D 06-MAR-2003.  
 X  
 F 15-AUG-2002; 2002WO-US026319.  
 X  
 X 16-AUG-2001; 2001US-0312726P.  
 R 21-DEC-2001; 2001US-00032376.  
 R 21-MAY-2002; 2002US-00153185.  
 X  
 X (KIMB ) KIMBERLY-CLARK WORLDWIDE INC.  
 X Quirk S, Weart IF;

DR WPI; 2003-381408/36.  
 XX Anti-angiogenic composition comprising peptide inhibitor of matrix  
 PT metalloproteinase, useful for decreasing the expression of vascular  
 PT endothelial growth factor and treating cancers and tissue injuries.  
 XX Claim 17; Page 15; 103pp; English.  
 PS  
 XX The present invention describes an anti-angiogenic composition (I) for  
 CC inhibiting expression of vascular endothelial growth factor (VEGF). (I)  
 CC comprises an effective amount of a peptide inhibitor of matrix  
 CC metalloproteinase (MMP), where the peptide can inhibit the expression of  
 CC VEGF. (I) has cytostatic, vulnary, cardiant, cerebroprotective,  
 CC antidiabetic, ophthalmological and dermatological activities. (I) can be  
 CC used for inhibiting expression of VEGF, and so can be used for inhibiting  
 CC growth of tumours and diminishing tumours size. The tumour can be  
 CC metastatic, non-metastatic, vascularised, non-vascularised, hard or soft.  
 CC (I) is also useful for treating injuries including wounds, surgical  
 CC incisions, chronic wounds, heart diseases and stroke. (I) is also useful  
 CC for treating disorders characterised by excessive angiogenesis e.g.  
 CC macular degeneration and diabetic retinopathy. The present sequence  
 CC represents a human MMP cleavage region peptide, which is used in the  
 CC exemplification of the present invention  
 XX Sequence 56 AA;  
 SQ

Query Match 95.0%; Score 19; DB 6; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 Db 31 VAEF 34

RESULT 60

ABG76312  
 ID ABG76312 standard; protein; 56 AA.  
 XX  
 AC ABG76312;  
 XX  
 DT 10-MAY-2003 (first entry)  
 XX Human matrix metalloproteinase (MMP) peptide inhibitor #4.  
 DE  
 XX Human; peptide inhibitor; matrix metalloproteinase-7; MMP-7;  
 KW cleavage region; proenzyme form; cellular proliferation; fibroblast;  
 KW keratinocyte; healthy skin development; wound healing; scarring;  
 KW skin tone; wrinkle; anti-aging; vulnary.  
 XX Homo sapiens.  
 OS  
 XX WO2003016520-A1.  
 PN  
 XX 27-FEB-2003.  
 PD  
 XX 15-AUG-2002; 2002WO-US026198.  
 PF  
 XX 16-AUG-2001; 2001US-0312726P.  
 PR 21-DEC-2001; 2001US-00032376.  
 PR 21-MAY-2002; 2002US-00153185.  
 XX  
 XX (KIMB ) KIMBERLY-CLARK WORLDWIDE INC.  
 PA  
 XX Quirk S, Malik S, Villanueva JM;  
 PI  
 XX WPI; 2003-289980/28.  
 DR  
 XX Novel peptide inhibitor of proteinase activity of matrix  
 PT metalloproteinases, e.g. matrix metalloproteinase-2, useful for  
 PT stimulating cellular proliferation of fibroblasts or keratinocytes.  
 X  
 X Claim 1; Page 16; 120pp; English.

XX The present invention relates to peptide inhibitors of metalloproteinases  
 CC (MMPs), particularly metalloproteinase-2 (MMP-2). The inhibitors have  
 CC peptide sequences related to the cleavage regions of the proenzyme forms  
 CC of the MMPs. The peptide inhibitors are useful for stimulating cellular  
 CC proliferation of fibroblasts or keratinocytes, promoting healthy skin  
 CC development, treating wounds, preventing scarring, improving skin tone,  
 CC reducing wrinkling and for stimulating the development of smooth, healthy  
 CC skin. The peptide inhibitors are useful as anti-aging and wound healing  
 CC compounds. ABG76309-ABG76321 represent peptide inhibitors of MMPs  
 XX  
 SQ Sequence 56 AA;

Query Match 95.0%; Score 19; DB 6; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAEF 5  
 ||||  
 DB 31 VAEF 34

RESULT 61  
 ID AAW55888 standard; protein; 58 AA.

XX AAW55888;  
 AC  
 XX  
 DT 22-JUL-1998 (first entry)  
 XX  
 DE Rat PC12.  
 XX  
 KW Rat; telomerase; human; cancer; screening; inhibitor; elucidation;  
 KW detection; probe; diagnosis; cell growth; ageing.  
 XX

OS Rattus sp.

XX WO9807838-A1.

XX 26-FEB-1998.

XX 21-AUG-1997; 97WO-JP002904.

XX 21-AUG-1996; 96JP-00219761.

PR 31-JAN-1997; 97JP-00018878.

PR 17-FEB-1997; 97JP-00031807.

XX (MITU) MITSUBISHI CHEM CORP.

XX Ishikawa F, Nakamura H, Takahashi K, Fujino Y, Harada N;

PI WPI; 1998-169149/15.

XX N-PSDB; AAV25990.

DR Telomerase protein of higher animals and humans and gene encoding it -  
 PT for use in diagnosis of cancer, screening of telomerase inhibitors and  
 PT elucidation of biological control mechanisms.

XX Example 1; Page 69-70; 106pp; Japanese.

XX The present sequence represents rat PC12 which is used in an example of  
 CC the present invention which describes protein components of telomerase.  
 CC The DNA or RNA encoding the telomerase protein component or its fragments  
 CC can be used as a nucleotide probe for the detection of cancer cells and  
 CC for diagnosis of cancer. Potential telomerase inhibitors can be screened  
 CC by measuring their effect on the assay of the active form in cells or  
 CC tissues. The polypeptide and DNA coding for it can be used in the  
 CC elucidation of biological control mechanisms of, e.g. cell growth or  
 CC ageing and of the mechanisms of cancer development

XX Sequence 58 AA;

Query Match 95.0%; Score 19; DB 2; Length 58;

Best Local Similarity 100.0%; Pred. No. 8.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAEF 5  
 ||||  
 DB 49 VAEF 52

RESULT 62  
 ID ABP04854 standard; protein; 58 AA.

XX ABP04854;

XX 24-JUN-2002 (first entry)

XX Human ORFX protein sequence SEQ ID NO:9690.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
 KW hypertension; hypothyroidism; cholesterol ester storage disease;  
 KW immune deficiency; immune disorder; infectious disease;  
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
 KW myasthenia gravis.

XX Homo sapiens.

XX WO200192523-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US010836.

XX 30-MAY-2000; 2000US-0206132P.

PR 29-AUG-2000; 2000US-0228716P.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach MD;

XX WPI; 2002-106308/14.

DR N-PSDB; ABN20606.

XX Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and autoimmune disorders.

XX Disclosure; SEQ ID NO 9690; 1037pp; English.

XX The present invention describes substantially purified human proteins  
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX  
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
 CC treating or preventing a pathology associated with an ORFX-associated  
 CC disorder in humans, and in the manufacture of a medicament for treating a  
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
 CC sequences can be used in gene therapy. ORFX sequences can be used in the  
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
 CC storage disease, various immune deficiencies and disorders, infectious  
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
 CC bone degenerative disorders, or periodontal disease, and for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues and conditions resulting from  
 CC systemic cytokine damage. N.B. The sequence data for this patent did not

C form part of the printed specification, but was obtained in electronic  
C format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
X Sequence 58 AA;  
Q Query Match 95.0%; Score 19; DB 5; Length 58;  
Best Local Similarity 100.0%; Pred. No. 8.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Y 2 VAEF 5  
b 30 VAEF 33  
ESULT 63  
AG60085  
D AAG60085 standard; protein; 59 AA.  
C AAG60085;  
T 18-OCT-2000 (first entry)  
X Arabidopsis thaliana protein fragment SEQ ID NO: 77787.  
E Arabidopsis thaliana.  
W Protein identification; signal transduction pathway; metabolic pathway;  
W hybridisation assay; genetic mapping; gene expression control; promoter;  
W termination sequence.  
X Arabidopsis thaliana.  
S EF1033405-A2.  
X 06-SEP-2000.  
D 25-FEB-2000; 2000EP-00301439.  
F 25-FEB-1999; 99US-0121825P.  
X 03-MAR-1999; 99US-0123180P.  
R 09-MAR-1999; 99US-0123548P.  
R 23-MAR-1999; 99US-0125788P.  
R 25-MAR-1999; 99US-0126264P.  
R 29-MAR-1999; 99US-0126785P.  
R 01-APR-1999; 99US-0127462P.  
R 06-APR-1999; 99US-0128234P.  
R 08-APR-1999; 99US-0128714P.  
R 16-APR-1999; 99US-0129845P.  
R 19-APR-1999; 99US-0130077P.  
R 21-APR-1999; 99US-0130449P.  
R 23-APR-1999; 99US-0130510P.  
R 28-APR-1999; 99US-0130891P.  
R 30-APR-1999; 99US-0131449P.  
R 30-APR-1999; 99US-0132048P.  
R 30-APR-1999; 99US-0132407P.  
R 04-MAY-1999; 99US-0132484P.  
R 05-MAY-1999; 99US-0132485P.  
R 06-MAY-1999; 99US-0132486P.  
R 06-MAY-1999; 99US-0132487P.  
R 07-MAY-1999; 99US-0132863P.  
R 11-MAY-1999; 99US-0134256P.  
R 14-MAY-1999; 99US-0134218P.  
R 14-MAY-1999; 99US-0134219P.  
R 14-MAY-1999; 99US-0134221P.  
R 14-MAY-1999; 99US-0134370P.  
R 18-MAY-1999; 99US-0134768P.  
R 19-MAY-1999; 99US-0134941P.  
R 20-MAY-1999; 99US-0135124P.  
R 21-MAY-1999; 99US-0135353P.  
R 24-MAY-1999; 99US-0135629P.  
R 25-MAY-1999; 99US-0136021P.  
R 27-MAY-1999; 99US-0136392P.  
R 28-MAY-1999; 99US-0136782P.  
R 01-JUN-1999; 99US-0137222P.  
R 03-JUN-1999; 99US-0137528P.  
R 04-JUN-1999; 99US-0137502P.  
R 07-JUN-1999; 99US-0137724P.  
R 08-JUN-1999; 99US-0138094P.  
R 10-JUN-1999; 99US-0138540P.  
R 10-JUN-1999; 99US-0138847P.  
R 14-JUN-1999; 99US-0139119P.  
R 16-JUN-1999; 99US-0139452P.  
R 16-JUN-1999; 99US-0139453P.  
R 17-JUN-1999; 99US-0139492P.  
R 18-JUN-1999; 99US-0139454P.  
R 18-JUN-1999; 99US-0139455P.  
R 18-JUN-1999; 99US-0139456P.  
R 18-JUN-1999; 99US-0139457P.  
R 18-JUN-1999; 99US-0139458P.  
R 18-JUN-1999; 99US-0139459P.  
R 18-JUN-1999; 99US-0139460P.  
R 18-JUN-1999; 99US-0139461P.  
R 18-JUN-1999; 99US-0139462P.  
R 18-JUN-1999; 99US-0139463P.  
R 18-JUN-1999; 99US-0139750P.  
R 18-JUN-1999; 99US-0139763P.  
R 21-JUN-1999; 99US-0139817P.  
R 22-JUN-1999; 99US-0139899P.  
R 23-JUN-1999; 99US-0140353P.  
R 23-JUN-1999; 99US-0140354P.  
R 24-JUN-1999; 99US-0140695P.  
R 28-JUN-1999; 99US-0140823P.  
R 29-JUN-1999; 99US-0140911P.  
R 30-JUN-1999; 99US-0141287P.  
R 01-JUL-1999; 99US-0141842P.  
R 01-JUL-1999; 99US-0142154P.  
R 02-JUL-1999; 99US-0142055P.  
R 06-JUL-1999; 99US-0142390P.  
R 08-JUL-1999; 99US-0142803P.  
R 09-JUL-1999; 99US-0142920P.  
R 12-JUL-1999; 99US-0142977P.  
R 13-JUL-1999; 99US-0143542P.  
R 14-JUL-1999; 99US-0143624P.  
R 15-JUL-1999; 99US-0144005P.  
R 16-JUL-1999; 99US-0144085P.  
R 16-JUL-1999; 99US-0144086P.  
R 19-JUL-1999; 99US-0144325P.  
R 19-JUL-1999; 99US-0144331P.  
R 19-JUL-1999; 99US-0144332P.  
R 19-JUL-1999; 99US-0144333P.  
R 19-JUL-1999; 99US-0144334P.  
R 19-JUL-1999; 99US-0144335P.  
R 20-JUL-1999; 99US-0144352P.  
R 20-JUL-1999; 99US-0144632P.  
R 21-JUL-1999; 99US-0144884P.  
R 21-JUL-1999; 99US-0144814P.  
R 21-JUL-1999; 99US-0145086P.  
R 21-JUL-1999; 99US-0145088P.  
R 22-JUL-1999; 99US-0145085P.  
R 22-JUL-1999; 99US-0145087P.  
R 22-JUL-1999; 99US-0145089P.  
R 22-JUL-1999; 99US-0145192P.  
R 23-JUL-1999; 99US-0145145P.  
R 23-JUL-1999; 99US-0145218P.  
R 23-JUL-1999; 99US-0145224P.  
R 26-JUL-1999; 99US-0145276P.  
R 27-JUL-1999; 99US-0145913P.  
R 27-JUL-1999; 99US-0145918P.  
R 27-JUL-1999; 99US-0145919P.  
R 28-JUL-1999; 99US-0145951P.  
R 02-AUG-1999; 99US-0146386P.  
R 02-AUG-1999; 99US-0146388P.  
R 02-AUG-1999; 99US-0146389P.  
R 04-AUG-1999; 99US-0147038P.  
R 04-AUG-1999; 99US-0147204P.  
R 04-AUG-1999; 99US-0147302P.  
R 05-AUG-1999; 99US-0147192P.  
R 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.  
PR 08-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 13-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 22-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159884P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161932P.  
PR 29-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 95.0%; Score 19; DB 3; Length 59;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 45 VAEF 48  
RESULT 64  
AAG60683  
ID AAG60683 standard; protein; 59 AA.  
XX AC AAG60683;  
XX 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 78631.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 23-APR-1999; 99US-0130891P.  
XX 28-APR-1999; 99US-0131449P.  
XX 30-APR-1999; 99US-0132048P.  
XX 30-APR-1999; 99US-0132407P.  
XX 04-MAY-1999; 99US-0132484P.  
XX 05-MAY-1999; 99US-0132485P.  
XX 06-MAY-1999; 99US-0132486P.  
XX 06-MAY-1999; 99US-0132487P.  
XX 07-MAY-1999; 99US-0132863P.  
XX 11-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134219P.  
XX 14-MAY-1999; 99US-0134221P.  
XX 14-MAY-1999; 99US-0134370P.  
XX 18-MAY-1999; 99US-0134768P.  
XX 19-MAY-1999; 99US-0134941P.  
XX 20-MAY-1999; 99US-0135124P.  
XX 21-MAY-1999; 99US-0135353P.  
XX 24-MAY-1999; 99US-0135629P.  
XX 25-MAY-1999; 99US-0136021P.  
XX 27-MAY-1999; 99US-0136392P.  
XX 28-MAY-1999; 99US-0136782P.  
XX 01-JUN-1999; 99US-0137222P.  
XX 03-JUN-1999; 99US-0137528P.  
XX 04-JUN-1999; 99US-0137502P.  
XX 07-JUN-1999; 99US-0137724P.  
XX 08-JUN-1999; 99US-0138094P.  
XX 10-JUN-1999; 99US-0138540P.  
XX 10-JUN-1999; 99US-0138847P.  
XX 14-JUN-1999; 99US-0139119P.  
XX 16-JUN-1999; 99US-0139452P.

R 16-JUN-1999; 99US-0139453P.  
R 17-JUN-1999; 99US-0139452P.  
R 18-JUN-1999; 99US-0139454P.  
R 18-JUN-1999; 99US-0139455P.  
R 18-JUN-1999; 99US-0139456P.  
R 18-JUN-1999; 99US-0139457P.  
R 18-JUN-1999; 99US-0139458P.  
R 18-JUN-1999; 99US-0139459P.  
R 18-JUN-1999; 99US-0139460P.  
R 18-JUN-1999; 99US-0139461P.  
R 18-JUN-1999; 99US-0139462P.  
R 18-JUN-1999; 99US-0139463P.  
R 18-JUN-1999; 99US-0139750P.  
R 18-JUN-1999; 99US-0139763P.  
R 21-JUN-1999; 99US-0139817P.  
R 22-JUN-1999; 99US-0139899P.  
R 23-JUN-1999; 99US-0140353P.  
R 23-JUN-1999; 99US-0140354P.  
R 24-JUN-1999; 99US-0140695P.  
R 28-JUN-1999; 99US-0140823P.  
R 28-JUN-1999; 99US-0140891P.  
R 30-JUN-1999; 99US-0141287P.  
R 01-JUL-1999; 99US-0141842P.  
R 01-JUL-1999; 99US-0142154P.  
R 06-JUL-1999; 99US-0142055P.  
R 06-JUL-1999; 99US-0142390P.  
R 08-JUL-1999; 99US-0142803P.  
R 09-JUL-1999; 99US-0142920P.  
R 12-JUL-1999; 99US-0142977P.  
R 13-JUL-1999; 99US-0143542P.  
R 14-JUL-1999; 99US-0143624P.  
R 15-JUL-1999; 99US-0144005P.  
R 15-JUL-1999; 99US-0144085P.  
R 16-JUL-1999; 99US-0144086P.  
R 19-JUL-1999; 99US-0144325P.  
R 19-JUL-1999; 99US-0144331P.  
R 19-JUL-1999; 99US-0144332P.  
R 19-JUL-1999; 99US-0144333P.  
R 19-JUL-1999; 99US-0144334P.  
R 19-JUL-1999; 99US-0144335P.  
R 20-JUL-1999; 99US-0144352P.  
R 20-JUL-1999; 99US-0144632P.  
R 20-JUL-1999; 99US-0144684P.  
R 21-JUL-1999; 99US-0144614P.  
R 21-JUL-1999; 99US-0145086P.  
R 21-JUL-1999; 99US-0145088P.  
R 21-JUL-1999; 99US-0145085P.  
R 22-JUL-1999; 99US-0145087P.  
R 22-JUL-1999; 99US-0145089P.  
R 22-JUL-1999; 99US-0145192P.  
R 23-JUL-1999; 99US-0145145P.  
R 23-JUL-1999; 99US-0145218P.  
R 23-JUL-1999; 99US-0145224P.  
R 25-JUL-1999; 99US-0145276P.  
R 27-JUL-1999; 99US-0145313P.  
R 27-JUL-1999; 99US-0145318P.  
R 27-JUL-1999; 99US-0145319P.  
R 28-JUL-1999; 99US-0145351P.  
R 02-AUG-1999; 99US-0146386P.  
R 02-AUG-1999; 99US-0146388P.  
R 02-AUG-1999; 99US-0146389P.  
R 03-AUG-1999; 99US-0147038P.  
R 04-AUG-1999; 99US-0147204P.  
R 04-AUG-1999; 99US-0147302P.  
R 05-AUG-1999; 99US-0147192P.  
R 05-AUG-1999; 99US-0147260P.  
R 06-AUG-1999; 99US-0147303P.  
R 06-AUG-1999; 99US-0147416P.  
R 09-AUG-1999; 99US-0147493P.  
R 09-AUG-1999; 99US-0147935P.  
R 10-AUG-1999; 99US-0148171P.  
R 11-AUG-1999; 99US-0148319P.  
R 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154033P.  
PR 22-SEP-1999; 99US-0154779P.  
PR 23-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 26-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157533P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 28-OCT-1999; 99US-0162142P.

Query Match 95.0%; Score 19; DB 3; Length 59;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 45 VAEF 48



## RESULT 65

ABB42044  
ID ABB42044 standard; peptide; 60 AA.

XX AC ABB42044;  
XX 04-FEB-2002 (first entry)  
DT  
DE Peptide #9550 encoded by human foetal liver single exon probe.  
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX OS Homo sapiens.

XX PN WO200157277-A2.  
XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000669.  
XX 04-FEB-2000; 2000US-0180312P.  
XX 26-MAY-2000; 2000US-0207456P.  
XX 30-JUN-2000; 2000US-00608408.  
XX 03-AUG-2000; 2000US-00632366.  
XX 21-SEP-2000; 2000US-0234687P.  
XX 27-SEP-2000; 2000US-0236359P.  
XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT Gene expression in human foetal liver.  
XX Claim 27; SEQ ID NO 34679; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring  
CC human gene expression in a sample derived from human foetal liver. The  
CC single exon nucleic acid probes may be used for predicting, measuring and  
CC displaying gene expression in samples derived from human foetal liver. The  
CC present sequence is a peptide encoded by a single exon nucleic acid probe  
CC of the invention. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 60 AA;  
SQ

Query Match 95.0%; Score 19; DB 4; Length 60;  
Best Local Similarity 100.0%; Pred. No. 8.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 8 VAEF 11

## RESULT 66

AAM35846  
ID AAM35846 standard; protein; 60 AA.

XX AC AAM35846;  
XX 17-OCT-2001 (first entry)  
DT  
DE Peptide #9883 encoded by probe for measuring placental gene expression.  
XX KW Probe; microarray; human; placenta; antenatal diagnosis;  
XX genetic disorder.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000663.

XX 04-FEB-2000; 2000US-0180312P.  
XX 26-MAY-2000; 2000US-0207456P.  
XX 30-JUN-2000; 2000US-00608408.  
XX 03-AUG-2000; 2000US-00632366.  
XX 21-SEP-2000; 2000US-0234687P.  
XX 27-SEP-2000; 2000US-0236359P.  
XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human placenta.

XX PS Claim 27; SEQ ID NO 36115; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP;  
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders

XX SQ Sequence 60 AA;  
SQ

Query Match 95.0%; Score 19; DB 4; Length 60;  
Best Local Similarity 100.0%; Pred. No. 8.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 8 VAEF 11

## RESULT 67

AAM75737  
ID AAM75737 standard; protein; 60 AA.

XX AC AAM75737;

XX 06-NOV-2001 (first entry)

XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36043.  
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;  
XX microarray; cancer; leukaemia; lymphoma; myeloma.

XX OS Homo sapiens.

XX PN WO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000668.

XX 04-FEB-2000; 2000US-0180312P.  
XX 26-MAY-2000; 2000US-0207456P.  
XX 30-JUN-2000; 2000US-00608408.  
XX 03-AUG-2000; 2000US-00632366.  
XX 21-SEP-2000; 2000US-0234687P.  
XX 27-SEP-2000; 2000US-0236359P.  
XX 04-OCT-2000; 2000GB-00024263.

XX OS

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human bone marrow.

XX Example 4; SEQ ID NO 36043; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is a

CC protein encoded by one of the probes of the invention

XX Sequence 60 AA;

Query Match 95.0%; Score 19; DB 4; Length 60;

Best Local Similarity 100.0%; Pred. No. 8.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

DB 8 VAEF 11

RESULT 68

AA62925

ID AAM62925 standard; protein; 60 AA.

AC AAM62925;

DT 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe encoded protein SEQ ID NO: 35030.

XX Human; brain expressed exon; gene expression analysis; probe; microarray;

KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

XX Homo sapiens.

OS WO200157275-A2.

PN 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000667.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains.

XX Example 4; SEQ ID NO 35030; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is a protein encoded by one of

XX the probes of the invention

XX Sequence 60 AA;

Query Match 95.0%; Score 19; DB 4; Length 60;

Best Local Similarity 100.0%; Pred. No. 8.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

DB 8 VAEF 11

RESULT 69

ABG57475

ID ABG57475 standard; peptide; 60 AA.

XX AC ABG57475;

XX 25-FEB-2003 (first entry)

XX Human liver peptide, SEQ ID NO 36123.

XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

KW hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000664.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488998/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human adult liver.

XX Claim 27; SEQ ID NO 36123; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for

CC measuring human gene expression in a sample derived from human adult

CC liver, comprising one of 13109 defined nucleotide sequences given in the

CC specification (or complements/fragments). The probe hybridizes at high

CC stringency to a nucleic acid molecule expressed in the human adult liver.

CC (I) may be used for predicting, measuring and displaying gene expression

CC in samples derived from human adult liver. The genes identified may be

CC involved in genetic liver diseases such as cirrhosis,

CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is

CC associated with coronary heart disease. ABG47348-ABG59930 represent human

CC liver single exon encoded peptides of the invention. Note: The sequence

CC information for this patent does not appear in the printed specification

CC but was obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 60 AA;

Query Match 95.0%; Score 19; DB 4; Length 60;

Best Local Similarity 100.0%; Pred. No. 8.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 8 VAEF 11

## RESULT 70

AAV17262

ID AAY17262 standard; peptide; 62 AA.

XX AC AAY17262;

XX DT 09-AUG-1999 (first entry)

XX DE HLH domain of drosophila E-spl m5.

XX KW ORF polypeptide; dermatomyositis; polymyositis; inclusion body myositis;  
KW sarcoid myopathy; AZT myopathy; myocardial infarction; ischaemia; ERK-6;  
KW extracellular signal-regulated kinase; reperfusion.

XX OS Drosophila melanogaster.

XX PN WO9927099-A1.

XX PD 03-JUN-1999.

XX PF 23-NOV-1998; 98WO-US025046.

XX PR 25-NOV-1997; 97US-0066848P.

XX PA (SUGEN-) SUGEN INC.

XX PI Ullrich A, Giot J;

XX DR WPI; 1999-357837/30.

XX PT Nucleic acid encoding ORF polypeptide.

XX PS Disclosure; Fig 2; 79pp; English.

XX CC The invention relates to a human ORF polypeptide, a substrate for  
CC extracellular signal-regulated kinase, ERK-6. Substances that modulate  
CC the activity of the ORF polypeptide can be used to treat diseases  
CC selected from dermatomyositis, polymyositis, inclusion body myositis,  
CC sarcoid myopathy, AZT myopathy, myocardial infarction, and ischaemia/  
CC reperfusion. The probes and antibodies can be used to detect the presence  
CC of ORF in a sample

XX SQ Sequence 62 AA;

Query Match 95.0%; Score 19; DB 2; Length 62;

Best Local Similarity 100.0%; Pred. No. 8.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

DB 29 VAEF 32

## RESULT 71

AAM91247

ID AAM91247 standard; protein; 62 AA.

XX AC AAM91247;

XX DT 07-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen SEQ ID NO:18840.

XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX KW cytostatic; gene therapy; vaccine; metastasis.

XX OS Homo sapiens.  
XX WO200157182-A2.  
XX PD 09-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US001354.  
XX PR 31-JAN-2000; 2000US-0179065P.  
XX PR 04-FEB-2000; 2000US-0180628P.  
XX PR 24-FEB-2000; 2000US-0184664P.  
XX PR 02-MAR-2000; 2000US-0186350P.  
XX PR 16-MAR-2000; 2000US-0189874P.  
XX PR 17-MAR-2000; 2000US-0190076P.  
XX PR 18-APR-2000; 2000US-0198123P.  
XX PR 19-MAY-2000; 2000US-0205515P.  
XX PR 07-JUN-2000; 2000US-0209467P.  
XX PR 28-JUN-2000; 2000US-0214886P.  
XX PR 30-JUN-2000; 2000US-0215135P.  
XX PR 07-JUL-2000; 2000US-0216647P.  
XX PR 07-JUL-2000; 2000US-0216880P.  
XX PR 11-JUL-2000; 2000US-0217487P.  
XX PR 14-JUL-2000; 2000US-0218230P.  
XX PR 26-JUL-2000; 2000US-0220963P.  
XX PR 26-JUL-2000; 2000US-0220964P.  
XX PR 14-AUG-2000; 2000US-0224518P.  
XX PR 14-AUG-2000; 2000US-0224519P.  
XX PR 14-AUG-2000; 2000US-0225213P.  
XX PR 14-AUG-2000; 2000US-0225214P.  
XX PR 14-AUG-2000; 2000US-0225266P.  
XX PR 14-AUG-2000; 2000US-0225267P.  
XX PR 14-AUG-2000; 2000US-0225268P.  
XX PR 14-AUG-2000; 2000US-0225270P.  
XX PR 14-AUG-2000; 2000US-0225447P.  
XX PR 14-AUG-2000; 2000US-0225757P.  
XX PR 14-AUG-2000; 2000US-0225758P.  
XX PR 18-AUG-2000; 2000US-0226279P.  
XX PR 22-AUG-2000; 2000US-0226581P.  
XX PR 22-AUG-2000; 2000US-0226588P.  
XX PR 22-AUG-2000; 2000US-0227182P.  
XX PR 23-AUG-2000; 2000US-0227009P.  
XX PR 30-AUG-2000; 2000US-0228924P.  
XX PR 01-SEP-2000; 2000US-0229287P.  
XX PR 01-SEP-2000; 2000US-0229343P.  
XX PR 01-SEP-2000; 2000US-0229344P.  
XX PR 01-SEP-2000; 2000US-0229345P.  
XX PR 05-SEP-2000; 2000US-0229509P.  
XX PR 05-SEP-2000; 2000US-0229513P.  
XX PR 06-SEP-2000; 2000US-0230437P.  
XX PR 06-SEP-2000; 2000US-0230438P.  
XX PR 08-SEP-2000; 2000US-0231242P.  
XX PR 08-SEP-2000; 2000US-0231243P.  
XX PR 08-SEP-2000; 2000US-0231244P.  
XX PR 08-SEP-2000; 2000US-0231413P.  
XX PR 08-SEP-2000; 2000US-0231414P.  
XX PR 08-SEP-2000; 2000US-0232080P.  
XX PR 08-SEP-2000; 2000US-0232081P.  
XX PR 12-SEP-2000; 2000US-0231968P.  
XX PR 14-SEP-2000; 2000US-0232397P.  
XX PR 14-SEP-2000; 2000US-0232398P.  
XX PR 14-SEP-2000; 2000US-0232399P.  
XX PR 14-SEP-2000; 2000US-0232400P.  
XX PR 14-SEP-2000; 2000US-0232401P.  
XX PR 14-SEP-2000; 2000US-0233063P.  
XX PR 14-SEP-2000; 2000US-0233064P.  
XX PR 21-SEP-2000; 2000US-0233065P.  
XX PR 21-SEP-2000; 2000US-0234223P.  
XX PR 25-SEP-2000; 2000US-0234274P.  
XX PR 25-SEP-2000; 2000US-0234997P.  
XX PR 25-SEP-2000; 2000US-0234998P.



PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
 PT or for stimulating an immune response specific for a P. acnes protein.

PS Claim 7; SEQ ID NO 30343; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
 XX encoding a Propionibacterium acnes protein. The invention also relates to  
 XX polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
 XX immunogenic fragments of P. acnes polypeptides. The invention  
 XX additionally encompasses expression vectors and host cells comprising a  
 XX polynucleotide of the invention; antibodies against polypeptides of the  
 XX invention; fusion proteins comprising a polypeptide of the invention; a  
 XX method for stimulating an immune response specific for a P. acnes  
 XX polypeptide and an isolated T cell population comprising T cells prepared  
 XX via this method; a vaccine composition (comprising P. acnes polypeptides,  
 XX polynucleotides, antibodies, fusion proteins, T cell populations, or  
 XX antigen-presenting cells that express the polypeptide); a method and kit  
 XX for detecting or determining the presence or absence of P. acnes in a  
 XX patient; and a method for inhibiting the development of P. acnes in a  
 XX patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
 XX proteins, T cell populations or antigen-presenting cells that express the  
 XX polypeptides are useful for diagnosing, preventing or treating acne  
 XX vulgaris, or for stimulating an immune response specific for a P. acnes  
 XX protein. The polynucleotides can also be used as probes or primers for  
 XX nucleic acid hybridisation. The vaccine composition is useful for the  
 XX stimulation of an immune response against P. acnes, or for treating acne,  
 XX and the kit is useful for performing a diagnostic assay. The present  
 XX sequence represents a specifically claimed P. acnes polypeptide which is  
 XX thought to contain an immunogenic region. Note: The sequence data for  
 XX this patent did not form part of the printed specification, but was  
 XX obtained in electronic format directly from WIPO at  
 XX ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 62 AA;

Query Match 95.0%; Score 19; DB 6; Length 62;  
 Best Local Similarity 100.0%; Pred. NO. 8.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 DB 8 VAEF 11

RESULT 73  
 AA009517  
 ID AA009517 standard; protein; 64 AA.

XX AA009517;

XX 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 23409.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 XX tissue growth factor; immunomodulatory; cancer; leukaemia;  
 XX nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.  
 DR N-PSDB; AA189448.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 PT and treating e.g. leukemia, inflammation and immune disorders.

PS Claim 20; SEQ ID NO 23409; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and  
 CC the encoded proteins (AA00010-AA013910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 64 AA;

Query Match 95.0%; Score 19; DB 4; Length 64;  
 Best Local Similarity 100.0%; Pred. NO. 9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 DB 54 VAEF 57

RESULT 74

AA179941  
 ID AA179941 standard; protein; 67 AA.

XX AA179941;

XX 08-DEC-1999 (first entry)

XX Human prostate cancer-associated protein 49.

XX Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy;  
 KW cancer; tissue specificity; human.

XX Homo sapiens.

XX DE19811194-A1.

XX 16-SEP-1999.

XX 10-MAR-1998; 98DE-01011194.

XX 10-MAR-1998; 98DE-01011194.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

XX WPI; 1999-519629/44.

XX N-PSDB; AA233499.

XX New nucleic acid expressed at high level in normal prostatic tissue and  
 PT encoded polypeptides, used to treat cancer and screen for therapeutic  
 PT agents.

XX Claim 22; 143; 194pp; German.

XX This invention describes novel nucleic acid sequences (A) that are  
 CC expressed at high level in normal prostatic tissue. Polypeptides (I)  
 CC encoded by (A) are used: (a) for identifying agents for treatment of



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

M protein - protein search, using sw model

Run on: May 24, 2004, 17:32:13 ; Search time 12.5 seconds  
(without alignments)  
38.477 Million cell updates/sec

Title: US-09-594-978A-1

Perfect score: 20

Sequence: 1 XVAEF 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

PIR 78.\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	95.0	20	2 S39049	cytotoxin-binding
2	19	95.0	33	2 G64625	hypothetical prote
3	19	95.0	48	2 E64618	hypothetical prote
4	19	95.0	52	2 A42375	hypothetical prote
5	19	95.0	54	2 E85745	unknown protein en
6	19	95.0	59	2 A69386	conserved hypothet
7	19	95.0	59	2 E71349	probable preprotet
8	19	95.0	63	2 JC4002	carbamoycin resist
9	19	95.0	63	2 G64007	hypothetical prote
10	19	95.0	64	2 I51350	transferrin - Atla
11	19	95.0	74	2 D97045	hypothetical prote
12	19	95.0	75	2 S07277	gene D protein - p
13	19	95.0	79	2 T18034	hypothetical prote
14	19	95.0	79	2 T27605	hypothetical prote
15	19	95.0	84	2 F83922	hypothetical prote
16	19	95.0	87	2 A38725	transferrin - shee
17	19	95.0	89	1 H64116	ribosomal protein
18	19	95.0	89	2 F66818	hypothetical prote
19	19	95.0	89	2 T15018	hypothetical prote
20	19	95.0	96	2 AC3053	hypothetical prote
21	19	95.0	97	2 AB1397	conserved hypothet
22	19	95.0	99	2 E83244	conserved hypothet
23	19	95.0	101	2 H90297	conserved hypothet
24	19	95.0	106	2 T10097	nifX protein - Met
25	19	95.0	108	2 G64490	hypothetical prote
26	19	95.0	109	2 E70949	probable regulator
27	19	95.0	111	2 T17582	hypothetical prote
28	19	95.0	112	2 B83431	type III export pr
29	19	95.0	114	2 A55872	kedarcidin [valida

30	19	95.0	114	2 D97414	hypothetical prote
31	19	95.0	115	2 D75218	hypothetical prote
32	19	95.0	116	1 IPAF	insulin precursor
33	19	95.0	118	2 T48682	hypothetical prote
34	19	95.0	119	2 F95327	hypothetical prote
35	19	95.0	122	2 F82861	conjugal transfer
36	19	95.0	125	2 B69129	gamma-carboxymuon
37	19	95.0	125	2 D70790	hypothetical prote
38	19	95.0	126	2 B81368	hypothetical prote
39	19	95.0	127	2 D70799	hypothetical prote
40	19	95.0	127	2 AB1732	hypothetical prote
41	19	95.0	128	2 A54797	ileal lipid-bindin
42	19	95.0	129	2 T02012	probable cathepsin
43	19	95.0	130	2 D72348	conserved hypothet
44	19	95.0	131	2 G75208	hypothetical prote
45	19	95.0	132	2 JQ2270	hypothetical 15.2K
46	19	95.0	132	2 AG1996	hypothetical prote
47	19	95.0	133	2 AH0137	conserved hypothet
48	19	95.0	133	2 H75623	conserved hypothet
49	19	95.0	133	2 A84221	riboflavin synthas
50	19	95.0	134	1 WMS015	ybgC protein - Esc
51	19	95.0	134	2 C90725	hypothetical prote
52	19	95.0	134	2 D85576	hypothetical prote
53	19	95.0	134	2 AD0592	conserved hypothet
54	19	95.0	136	2 S76416	hypothetical prote
55	19	95.0	137	2 E81355	hypothetical prote
56	19	95.0	138	2 F82106	flagellar basal-bo
57	19	95.0	139	2 A69052	riboflavin synthas
58	19	95.0	139	2 A83975	hypothetical prote
59	19	95.0	140	2 B63323	hypothetical prote
60	19	95.0	142	2 JQ2279	low-temperature re
61	19	95.0	142	2 JQ2280	low-temperature re
62	19	95.0	143	2 H69515	riboflavin synthas
63	19	95.0	143	2 T36978	probable transposa
64	19	95.0	145	1 B41715	ribosomal protein
65	19	95.0	145	2 S77043	hypothetical prote
66	19	95.0	146	2 F89837	conserved hypothet
67	19	95.0	147	2 C84294	hypothetical prote
68	19	95.0	147	2 T34803	hypothetical prote
69	19	95.0	148	2 E81788	conserved hypothet
70	19	95.0	148	2 S10655	hypothetical prote
71	19	95.0	149	2 AC1137	ribose 5-phosphate
72	19	95.0	150	2 AB2857	pseudazurin limpo
73	19	95.0	150	2 AB1089	protein gp35 from
74	19	95.0	150	2 A11452	hypothetical prote
75	19	95.0	151	2 H87263	superoxide dismuta
76	19	95.0	152	2 S04623	hypothetical prote
77	19	95.0	154	2 B60184	hypothetical prote
78	19	95.0	155	2 I64130	PAL cross-reacting
79	19	95.0	156	2 G81283	ankyrin-repeat con
80	19	95.0	159	2 T17848	hypothetical prote
81	19	95.0	161	2 AE0981	probable acetylra
82	19	95.0	163	2 A24504	salivary glue prot
83	19	95.0	163	2 G70524	hypothetical prote
84	19	95.0	164	2 AD1229	B. subtilis yslB p
85	19	95.0	164	2 A69293	hypothetical prote
86	19	95.0	166	2 S05059	hypothetical prote
87	19	95.0	168	2 S48866	hypothetical 19.5K
88	19	95.0	168	2 AE0192	probable C4-dicarb
89	19	95.0	169	2 T36387	probable acetylra
90	19	95.0	169	2 AB2739	acetyltransferase
91	19	95.0	170	2 AC2917	invasion protein A
92	19	95.0	172	2 B75041	hypothetical prote
93	19	95.0	172	2 T49449	hemopoietic-specif
94	19	95.0	173	2 T22350	hypothetical prote
95	19	95.0	174	2 D70879	hypothetical prote
96	19	95.0	175	1 G71120	hypothetical prote
97	19	95.0	175	2 AC3648	conserved hypothet
98	19	95.0	175	2 I39055	Bcl-2 related - hu
99	19	95.0	176	2 C27805	ferritin chain M -
100	19	95.0	178	1 E63389	hypothetical prote
101	19	95.0	178	2 A38593	transcription fact
102	19	95.0	178	2 S03629	neurogenic gene co

103	19	95.0	179	2	S34345	hypothetical prote	176	19	95.0	219	2	H69199	conserved hypothet
104	19	95.0	180	2	D87031	probable acetyltra	177	19	95.0	219	2	AE1506	weakly carboxylest
105	19	95.0	180	2	E83558	hypothetical prote	178	19	95.0	220	2	TI0279	protein tyrosine p
106	19	95.0	180	2	A10676	probable exported	179	19	95.0	220	2	B84281	riboflavin-specifi
107	19	95.0	181	2	AG2202	cobinamide kinase	180	19	95.0	221	2	C83947	hypothetical prote
108	19	95.0	183	2	S13186	plasma retinol-bin	181	19	95.0	223	2	D69311	conserved hypothet
109	19	95.0	184	2	A90903	probable ante-term	182	19	95.0	223	2	T35665	hypothetical prote
110	19	95.0	184	2	F90873	probable anti-termi	183	19	95.0	223	2	AC0890	probable lipoprote
111	19	95.0	184	2	D71933	hypothetical prote	184	19	95.0	224	2	S69643	hypothetical prote
112	19	95.0	185	2	S43729	H+-transporting tw	185	19	95.0	224	2	AE1187	hypothetical prote
113	19	95.0	185	2	S39317	replication initia	186	19	95.0	226	2	AG3528	hypothetical prote
114	19	95.0	185	2	S39318	replication initia	187	19	95.0	228	2	A75290	chloramphenicol ac
115	19	95.0	185	2	D75635	resolvase - Deinoc	188	19	95.0	228	2	S48116	integral membrane
116	19	95.0	185	2	B97430	hypothetical prote	189	19	95.0	228	2	AD0693	pyridoxal kinase (
117	19	95.0	187	2	S16314	photosynthetic rea	190	19	95.0	228	2	D30857	hypothetical prote
118	19	95.0	187	2	AG2721	conserved hypothet	191	19	95.0	228	2	TI0399	hypothetical prote
119	19	95.0	187	2	AS5896	conserved hypothet	192	19	95.0	228	2	AH0081	probable exported
120	19	95.0	188	2	AB0132	conserved hypothet	193	19	95.0	229	2	F36721	probable glutathio
121	19	95.0	188	2	A71286	conserved hypothet	194	19	95.0	229	2	E70340	glutaredoxin-like
122	19	95.0	188	2	H85713	unknown protein en	195	19	95.0	230	2	E37440	DNA-3-methyladenin
123	19	95.0	189	2	T29159	hypothetical prote	196	19	95.0	230	2	S73229	ribosomal protein
124	19	95.0	189	2	H97633	pseudouridin precu	197	19	95.0	231	2	AF1144	phosphoglycerate m
125	19	95.0	190	2	A29413	ubiquinol-cytochro	198	19	95.0	231	2	AE1503	phosphoglycerate m
126	19	95.0	191	2	A82392	transcription regu	199	19	95.0	231	2	C83179	conserved hypothet
127	19	95.0	192	2	D97503	hypothetical prote	200	19	95.0	231	2	AF3380	hypothetical prote
128	19	95.0	193	1	C64328	conserved hypothet	201	19	95.0	232	2	B70653	hypothetical prote
129	19	95.0	194	1	RPECR5	resolvase - Escher	202	19	95.0	232	2	E87520	hypothetical prote
130	19	95.0	194	2	C84033	hypothetical prote	203	19	95.0	234	2	F96620	hypothetical prote
131	19	95.0	194	2	F97691	hypothetical prote	204	19	95.0	234	2	AC2185	hypothetical prote
132	19	95.0	194	2	B87453	RNA polymerase sig	205	19	95.0	234	2	E91119	hypothetical prote
133	19	95.0	195	2	B86204	hypothetical prote	206	19	95.0	234	2	D85964	hypothetical prote
134	19	95.0	196	1	BVECAU	transcription regu	207	19	95.0	234	2	S23360	hypothetical prote
135	19	95.0	196	2	A41853	hexose phosphate t	208	19	95.0	235	2	B97753	(p)ppGpp 3-pyropho
136	19	95.0	196	2	AH0963	Two-component syst	209	19	95.0	235	2	C95848	probable transcrip
137	19	95.0	196	2	F91204	transcription regu	210	19	95.0	236	2	AE9351	agropine synthesis
138	19	95.0	196	2	H86050	hypothetical prote	211	19	95.0	236	2	AE2931	agropine synthesis
139	19	95.0	196	2	A64888	probable resolvase	212	19	95.0	236	2	B70728	hypothetical prote
140	19	95.0	196	2	D64909	probable resolvase	213	19	95.0	236	2	A44686	hypothetical prote
141	19	95.0	196	2	T21347	hypothetical prote	214	19	95.0	236	2	TS0908	hypothetical prote
142	19	95.0	197	2	C86748	hypothetical prote	215	19	95.0	237	2	F83639	two-component resp
143	19	95.0	198	2	S48749	cytochrome P460 pr	216	19	95.0	237	2	C97085	hypothetical prote
144	19	95.0	199	1	VAHU	plasma retinol-bin	217	19	95.0	238	2	C86656	glycerol uptake fa
145	19	95.0	199	2	H90206	amidotransferase h	218	19	95.0	239	2	D81218	conserved hypothet
146	19	95.0	200	2	AB0620	probable bacteriop	219	19	95.0	239	2	F81795	probable periplasm
147	19	95.0	201	1	VA8B	retinol-binding pr	220	19	95.0	239	2	H75195	hypothetical prote
148	19	95.0	201	2	A39486	plasma retinol-bin	221	19	95.0	239	2	S31033	gene 88 protein -
149	19	95.0	201	2	I46257	retinol binding pr	222	19	95.0	240	2	D89281	sulfate ABC transp
150	19	95.0	201	2	T00799	hypothetical prote	223	19	95.0	240	2	B46066	probable ATP synth
151	19	95.0	202	2	T35126	hypothetical prote	224	19	95.0	241	2	TI4808	hypothetical prote
152	19	95.0	202	2	S77556	hypothetical prote	225	19	95.0	242	2	B70366	hypothetical prote
153	19	95.0	202	2	AH1169	hypothetical prote	226	19	95.0	242	2	G97959	hypothetical prote
154	19	95.0	204	2	C97866	hypothetical prote	227	19	95.0	247	2	S43728	H+-transporting tw
155	19	95.0	204	2	G89753	protein FlilC7.1 [l	228	19	95.0	248	2	AF0177	probable beta-keto
156	19	95.0	205	2	S30739	hypothetical prote	229	19	95.0	248	2	TI2632	water channel prot
157	19	95.0	205	2	H91225	hypothetical prote	230	19	95.0	248	2	TI2632	probable tonoplast
158	19	95.0	205	2	F86072	hypothetical prote	231	19	95.0	248	2	F72604	probable high-affi
159	19	95.0	205	2	G01942	mitotic feedback c	232	19	95.0	248	2	TI8315	hypothetical prote
160	19	95.0	208	2	A48642	hypothetical prote	233	19	95.0	249	2	G81693	serine/threonine p
161	19	95.0	208	2	AH2658	3-methyladenine-DN	234	19	95.0	249	2	B69343	conserved hypothet
162	19	95.0	209	2	H97519	hypothetical prote	235	19	95.0	249	2	TI9088	hypothetical prote
163	19	95.0	212	2	A83521	conserved hypothet	236	19	95.0	250	2	UJ0102	TobR57-18C protein
164	19	95.0	213	2	G82067	hypothetical prote	237	19	95.0	250	2	S13719	probable membrane
165	19	95.0	213	2	AF2551	hypothetical prote	238	19	95.0	250	2	S1781	integral membrane
166	19	95.0	213	2	A84248	transcription regu	239	19	95.0	251	2	D64580	hypothetical prote
167	19	95.0	214	2	T30310	probable GTP bindi	240	19	95.0	251	2	F69494	(R)-hydroxyglutary
168	19	95.0	215	2	A13072	transcription regu	241	19	95.0	252	2	C90422	hypothetical prote
169	19	95.0	215	2	F84889	hypothetical prote	242	19	95.0	252	2	F87575	hypothetical prote
170	19	95.0	218	2	H83485	conserved hypothet	243	19	95.0	252	2	G44020	hypothetical prote
171	19	95.0	218	2	T03287	osmotin protein ho	244	19	95.0	253	1	S17711	probable dehydroge
172	19	95.0	218	2	D71392	coat protein - cuc	245	19	95.0	253	2	C65067	2-deoxy-D-gluconat
173	19	95.0	218	2	AB0254	probable fumarylac	246	19	95.0	253	2	C91091	2-deoxy-D-gluconat
174	19	95.0	218	2	T47889	hypothetical prote	247	19	95.0	253	2	F85936	2-deoxy-D-gluconat
175	19	95.0	219	2	T35128	hypothetical prote	248	19	95.0	253	2	AD0210	2-deoxy-D-gluconat



249	19	95.0	233	2	B72552	probable ATP-depen	322	19	95.0	291	2	G59479	methionyl aminopep
250	19	95.0	234	2	F90087	ATP-dependent Clp	323	19	95.0	291	2	E71491	probable geranyl t
251	19	95.0	254	2	T02078	phocysystem II oxy	324	19	95.0	291	2	H96971	N-terminal Chey re
252	19	95.0	254	2	AB0161	deoxyribose operon	325	19	95.0	292	1	C64503	conserved hypoteth
253	19	95.0	255	2	S12985	phosphoprotein phi	326	19	95.0	292	2	T32000	hypothetical prote
254	19	95.0	256	2	T52130	probable copper/zi	327	19	95.0	293	2	F71344	hypothetical prote
255	19	95.0	257	2	T35273	probable enoyl coA	328	19	95.0	294	2	H64465	methionyl aminopep
256	19	95.0	257	2	S56108	26S proteasome reg	329	19	95.0	294	2	A48934	nucleic acid-bindi
257	19	95.0	258	2	B83275	conserved hypoteth	330	19	95.0	295	2	D82341	ribosomal protein
258	19	95.0	258	2	F69479	nicotinate-nucleot	331	19	95.0	295	2	A49906	glucose-1-phosphat
259	19	95.0	258	2	T36359	hypothetical prote	332	19	95.0	295	2	C82828	glucose-1-phosphat
260	19	95.0	259	2	S75599	hypothetical prote	333	19	95.0	295	2	C43718	hypothetical prote
261	19	95.0	259	2	G81427	periplasmic protei	334	19	95.0	297	2	A84767	hypothetical prote
262	19	95.0	260	2	AF2054	hypothetical prote	335	19	95.0	297	2	F64470	sulfate permease (
263	19	95.0	261	2	AD2895	short-chain dehydr	336	19	95.0	297	2	C20002	transposase alr156
264	19	95.0	261	2	G97670	probable oxidoredu	337	19	95.0	297	2	AD1931	transposase alr099
265	19	95.0	261	2	T00100	hypothetical prote	338	19	95.0	297	2	AD2141	transposase alr268
266	19	95.0	261	2	G84010	hypothetical prote	339	19	95.0	297	2	AG2142	transposase alr269
267	19	95.0	261	2	H98213	transcription repr	340	19	95.0	297	2	AI2138	transposase alr366
268	19	95.0	262	2	B72333	bacteriocin - Ther	341	19	95.0	297	2	AI2228	transposase alr338
269	19	95.0	267	2	AC0132	probable aldo/keto	342	19	95.0	297	2	AI2258	transposase all362
270	19	95.0	267	2	A57490	matrilysin (EC 3.4	343	19	95.0	298	2	H98170	myo-inositol catab
271	19	95.0	267	2	T24397	hypothetical prote	344	19	95.0	298	2	AE3116	myo-inositol catab
272	19	95.0	267	2	A87404	ribosomal protein	345	19	95.0	299	2	G82393	transcription regu
273	19	95.0	267	2	AI3121	ABC transporter, m	346	19	95.0	300	2	JC4367	NAD(P)-arginine AD
274	19	95.0	267	2	F98165	hypothetical prote	347	19	95.0	301	2	JC5811	osteopontin - rat
275	19	95.0	269	1	R5BYL3	ribosomal protein	348	19	95.0	302	2	A96789	cytochrome-c perox
276	19	95.0	269	2	B44499	major intrinsic pr	349	19	95.0	304	2	G81417	hypothetical prote
277	19	95.0	269	2	A41616	erythrocyte integr	350	19	95.0	305	2	S40927	hypothetical prote
278	19	95.0	269	2	JC1320	water channel prot	351	19	95.0	305	2	G71441	hypothetical prote
279	19	95.0	269	2	I53266	uterine water chan	352	19	95.0	306	2	B98164	hypothetical prote
280	19	95.0	269	2	C71228	hypothetical prote	353	19	95.0	306	2	AE3123	glycosyltransferas
281	19	95.0	269	2	A56487	signal recognition	354	19	95.0	307	2	T19582	hypothetical prote
282	19	95.0	271	2	JC2348	water channel prot	355	19	95.0	308	2	D86742	thioredoxin-disulf
283	19	95.0	272	2	B89377	conserved hypoteth	356	19	95.0	309	1	S62835	dnaj protein homol
284	19	95.0	274	2	S37705	amvC protein - The	357	19	95.0	309	2	D90520	conserved hypoteth
285	19	95.0	274	2	H88690	protein p41H10.8 [	358	19	95.0	309	2	S76393	hypothetical prote
286	19	95.0	274	2	C83709	hypothetical prote	359	19	95.0	309	2	T09564	glutaminyl-peptide
287	19	95.0	274	2	H97337	acetyl-CoA carboxy	360	19	95.0	310	2	T03596	phosphoprotein pho
288	19	95.0	275	2	B97260	dihydroneopterin a	361	19	95.0	310	2	C90245	conserved hypoteth
289	19	95.0	276	2	AG1842	uroporphyrinogen-I	362	19	95.0	311	2	A95117	dihydrodipicolinat
290	19	95.0	276	2	T19845	hypothetical prote	363	19	95.0	311	2	G97986	dihydrodipicolinat
291	19	95.0	277	2	JW0102	azaarene carbazole	364	19	95.0	311	2	A99196	purine nucleosidas
292	19	95.0	277	2	A72420	conserved hypoteth	365	19	95.0	311	2	B84341	hypothetical prote
293	19	95.0	278	2	E84983	hypothetical prote	366	19	95.0	312	2	S24264	phosphoprotein pho
294	19	95.0	278	2	A81040	conserved hypoteth	367	19	95.0	312	2	S31086	phosphoprotein pho
295	19	95.0	278	2	E71293	probable nicotinam	368	19	95.0	312	2	S31089	phosphoprotein pho
296	19	95.0	279	2	S52582	prephenate dehydra	369	19	95.0	312	2	B23705	cysteine proteinas
297	19	95.0	279	2	F86842	prephenate dehydra	370	19	95.0	312	2	S67052	hypothetical prote
298	19	95.0	279	2	T22051	hypothetical prote	371	19	95.0	313	2	A26485	dnak-type molecula
299	19	95.0	279	2	D84185	hypothetical prote	372	19	95.0	313	2	S27698	phosphoenolpyruvat
300	19	95.0	280	2	S71506	site-specific DNA-	373	19	95.0	313	2	E45509	desiccation-relate
301	19	95.0	280	2	AC3579	transcription regu	374	19	95.0	313	2	E95853	hypothetical prote
302	19	95.0	281	2	C86221	hypothetical prote	375	19	95.0	313	2	AD1285	glycerate dehydrog
303	19	95.0	281	2	E84637	hypothetical prote	376	19	95.0	313	2	AG1656	phosphoprotein pho
304	19	95.0	282	2	F97252	fof1-type ATP synt	377	19	95.0	314	2	S52371	phosphoprotein pho
305	19	95.0	282	2	B70200	trNA-pseudouridine	378	19	95.0	314	2	H69458	potassium channel
306	19	95.0	282	2	H83651	hypothetical prote	379	19	95.0	314	2	AE2517	hypothetical prote
307	19	95.0	283	2	T31275	2-hydroxyuconate-	380	19	95.0	316	2	S26225	phosphoprotein pho
308	19	95.0	284	2	T37996	probable aldose re	381	19	95.0	316	2	S42397	DNA-(apurinic or a
309	19	95.0	284	2	B70703	hypothetical prote	382	19	95.0	316	2	T32993	hypothetical prote
310	19	95.0	285	2	A84494	envelope-like prot	383	19	95.0	317	1	A25917	osteopontin precu
311	19	95.0	285	2	JC5419	2-hydroxyuconate-	384	19	95.0	317	2	T03594	phosphoprotein pho
312	19	95.0	286	1	AI0288	pyridoxal kinase (	385	19	95.0	317	2	A39500	DNA-(apurinic or a
313	19	95.0	287	2	S07563	glutamine-tRNA lig	386	19	95.0	317	2	F82672	ATP sulfurylase, s
314	19	95.0	287	2	F64920	probable pyridoxal	387	19	95.0	317	2	D69258	conserved hypoteth
315	19	95.0	287	2	A90922	pyridoxal kinase 2	388	19	95.0	318	2	S20882	phosphoprotein pho
316	19	95.0	287	2	E85770	pyridoxal kinase 2	389	19	95.0	318	2	T13015	phosphoprotein pho
317	19	95.0	287	2	T35229	hypothetical prote	390	19	95.0	318	2	S23550	DNA-(apurinic or a
318	19	95.0	288	2	A13400	glycine hydroxymet	391	19	95.0	319	2	T05515	ubiquitin-activati
319	19	95.0	289	1	G64167	hypothetical prote	392	19	95.0	320	2	D86259	protein T12C24.6 [
320	19	95.0	290	2	AD2870	ABC transporter, m	393	19	95.0	320	2	D96750	unknown protein F2
321	19	95.0	290	2	F97646	probable aliphatic	394	19	95.0	320	2	G96834	probable oxidoredu

395	19	95.0	320	2	T28992	hypotheical prote	468	19	95.0	352	2	AC1560	two-component sens
396	19	95.0	322	2	D83785	hypotheical prote	469	19	95.0	352	2	AE1202	two-component sens
397	19	95.0	323	2	T09550	phosphoprotein pho	470	19	95.0	353	2	D97295	parvulin-like pept
398	19	95.0	323	2	G84325	hypotheical prote	471	19	95.0	355	2	AG1877	hypotheical prote
399	19	95.0	324	2	E84505	hypotheical prote	472	19	95.0	355	2	AI3624	spemidine/putresc
400	19	95.0	325	2	JC7553	brain mitochondria	473	19	95.0	357	2	AC1276	aminopeptidase hom
401	19	95.0	326	2	G69366	homoserine dehydro	474	19	95.0	357	2	AC1639	aminopeptidase hom
402	19	95.0	326	2	T05094	peroxidase homolog	475	19	95.0	357	2	A84551	hypotheical prote
403	19	95.0	326	2	T19070	hypotheical prote	476	19	95.0	359	2	T02011	probable cathepsin
404	19	95.0	327	2	AC3482	sulfate-binding pr	477	19	95.0	359	2	D95316	probable ABC-trans
405	19	95.0	327	2	E95968	hypotheical expor	478	19	95.0	359	2	H95382	probable ABC-trans
406	19	95.0	328	2	E75564	ABC transporter, A	479	19	95.0	360	2	D69025	pleiotropic regula
407	19	95.0	328	2	A89867	hypotheical prote	480	19	95.0	360	2	AB3078	hypotheical prote
408	19	95.0	328	2	C70368	conserved hypotet	481	19	95.0	361	2	G98208	probable ATP-bind
409	19	95.0	328	2	H84548	hypotheical prote	482	19	95.0	361	2	AH3042	hypotheical prote
410	19	95.0	329	2	AD2629	33 kda chaperonin,	483	19	95.0	361	2	D98243	cpuaa (AF234619), l
411	19	95.0	329	2	T50117	mo25 homolog [impo	484	19	95.0	361	2	AE1979	ABC transporter AT
412	19	95.0	329	2	T29218	hypotheical prote	485	19	95.0	361	2	E96743	probable deoxyguan
413	19	95.0	330	1	H64404	hypotheical prote	486	19	95.0	362	2	C97448	hypotheical prote
414	19	95.0	330	2	A83417	probable oxidoredu	487	19	95.0	363	2	E95371	probable oxidoredu
415	19	95.0	331	2	T35499	probable D-lactate	488	19	95.0	363	2	A81134	acyl-CoA dehydrog
416	19	95.0	331	2	AE2928	dehydrogenase Atu3	489	19	95.0	363	2	C81888	probable acyl-CoA
417	19	95.0	331	2	D82972	probable lipolytic	490	19	95.0	364	2	G70364	conserved hypotet
418	19	95.0	331	2	AE2666	divalent cation tr	491	19	95.0	365	2	S76544	hypotheical prote
419	19	95.0	333	2	D75067	probable dehydrog	492	19	95.0	366	1	G49964	photosynthetic rea
420	19	95.0	333	2	C71165	probable dehydrog	493	19	95.0	366	2	T50891	cytochrome subunit
421	19	95.0	333	2	A41881	collagenase Prc (	494	19	95.0	366	2	G84249	NADH-dependent fla
422	19	95.0	334	1	S15318	transcription regu	495	19	95.0	368	2	G70462	succinate-CoA liga
423	19	95.0	334	2	T49195	hypotheical prote	496	19	95.0	368	2	G97291	hypotheical prote
424	19	95.0	334	2	A95951	hypotheical prote	497	19	95.0	368	2	S75652	ABC-type transport
425	19	95.0	335	2	S42807	HSR203J protein -	498	19	95.0	368	2	C83898	response regulator
426	19	95.0	336	2	B95906	probable dehydrog	499	19	95.0	368	2	A97189	lps biosynthesis p
427	19	95.0	336	2	AE0337	probable aspartate	500	19	95.0	369	2	S56638	mitogen-activated
428	19	95.0	336	2	AH3569	ABC transporter AT	501	19	95.0	369	2	A95124	RNA polymerase sig
429	19	95.0	336	2	C87623	cobalamin biosynth	502	19	95.0	369	2	C97994	RNA polymerase sig
430	19	95.0	336	2	T00571	dolichyl-phosphate	503	19	95.0	369	2	B83571	probable ATP-bind
431	19	95.0	337	2	T35862	probable secreted	504	19	95.0	369	2	S77028	protein kinase, 41
432	19	95.0	337	2	T27857	hypotheical prote	505	19	95.0	369	2	T51477	glutamine-rich pro
433	19	95.0	338	2	AG2938	2-hydroxyacid-fam	506	19	95.0	370	2	F36819	C14 protein - rabo
434	19	95.0	338	2	H98343	hypotheical prote	507	19	95.0	370	2	F95927	probable mandelate
435	19	95.0	338	2	E75301	conserved hypotet	508	19	95.0	370	2	C97998	platelet-derived g
436	19	95.0	339	2	S17930	transcription init	509	19	95.0	371	2	T09622	protein kinase MMK
437	19	95.0	339	2	JC7509	glycoprotein VI-1	510	19	95.0	371	2	F86625	GRP-binding protei
438	19	95.0	339	2	T25630	hypotheical prote	511	19	95.0	371	2	D95000	GRP-binding protei
439	19	95.0	340	2	T26125	phosphotransferase	512	19	95.0	371	2	D97872	conserved hypotet
440	19	95.0	340	2	T32646	hypotheical prote	513	19	95.0	371	2	AH0686	hypotheical prote
441	19	95.0	340	2	E83146	membrane-bound lyt	514	19	95.0	372	2	AD0075	probable transport
442	19	95.0	340	2	I49451	alpha 4 protein -	515	19	95.0	373	2	AC0253	ribonuclease iii (
443	19	95.0	341	2	S71223	xyloglucan endo-1,	516	19	95.0	375	2	AB3101	dipeptidase [impor
444	19	95.0	342	2	AH0087	sigma-54 transcrip	517	19	95.0	375	2	H98185	membrane dipeptida
445	19	95.0	342	2	S29894	strictosidine synt	518	19	95.0	375	2	T16116	hypotheical prote
446	19	95.0	342	2	AB2258	cobS protein [impo	519	19	95.0	377	2	A48352	genome polypeptin
447	19	95.0	343	2	B86446	probable endoxylg	520	19	95.0	377	2	D95861	probable ABC trans
448	19	95.0	343	2	C88986	protein C50H11.7 (	521	19	95.0	378	1	B38178	cella protein pla
449	19	95.0	343	2	D84390	sulfate transport	522	19	95.0	379	2	AF2409	mannosyl transfera
450	19	95.0	344	1	RWRTC2	T-cell surface gly	523	19	95.0	379	2	H82284	queuine tRNA-ribos
451	19	95.0	344	2	A98354	hypotheical prote	524	19	95.0	379	2	G97201	NAD(PAD)-depend
452	19	95.0	344	2	B28967	T-cell surface gly	525	19	95.0	379	2	JC7710	NADH-rubredoxin ox
453	19	95.0	344	2	I49582	CD2 antigen protei	526	19	95.0	379	2	AD3390	pbnM protein [impo
454	19	95.0	344	2	AI1902	4-hydroxyphenylpyr	527	19	95.0	379	2	BE3325	ABC transporter, m
455	19	95.0	344	2	S01325	strictosidine synt	528	19	95.0	380	2	D64533	cystathionine gamm
456	19	95.0	345	2	F02039	anthranilate phosp	529	19	95.0	380	2	D71973	probable cystathio
457	19	95.0	345	2	E83927	hypotheical prote	530	19	95.0	381	2	S46583	442K curved dna-bi
458	19	95.0	346	2	A45885	MHC class I histoc	531	19	95.0	383	2	AG3304	pleiotropic regula
459	19	95.0	346	2	C95404	hypotheical prote	532	19	95.0	384	2	G72777	probable S2p metal
460	19	95.0	346	2	C95090	hypotheical prote	533	19	95.0	384	2	D86821	hydroxymethylgluta
461	19	95.0	347	2	D69373	immunogenic protei	534	19	95.0	385	2	T16447	hypotheical prote
462	19	95.0	347	2	AF2645	flagellar motor sw	535	19	95.0	385	2	H84411	phosphoglycerate k
463	19	95.0	347	2	F97427	flagellar motor sw	536	19	95.0	386	2	D64074	hypotheical prote
464	19	95.0	348	2	E69162	sulfate transport	537	19	95.0	386	2	B97411	33k chaperonin (he
465	19	95.0	349	2	H72714	probable O-sialogl	538	19	95.0	387	2	AC3190	hypotheical prote
466	19	95.0	351	2	D84541	hypotheical prote	539	19	95.0	388	2	E70894	probable metB prot
467	19	95.0	352	2	S22464	strictosidine synt	540	19	95.0	388	2	F87208	cystathionine [gam

541	19	95.0	388	2	T23670	hypothetical prote	614	19	95.0	427	2	G98259	methionine gamma-l
542	19	95.0	389	2	T15102	hypothetical prote	615	19	95.0	428	2	F69050	glycine hydroxymet
543	19	95.0	390	2	G84245	NADH dehydrogenase	616	19	95.0	429	2	G72323	phosphopyruvate hy
544	19	95.0	390	2	E69272	conserved hypothet	617	19	95.0	429	2	S30049	transcription fact
545	19	95.0	390	2	H75023	nicotinate phospho	618	19	95.0	429	2	AC2227	hypothetical prote
546	19	95.0	390	2	A84226	farnesyl-diphospha	619	19	95.0	430	2	AC3975	O-acetylhomoserine
547	19	95.0	390	2	T25996	hypothetical prote	620	19	95.0	430	2	B95892	probable ABC trans
548	19	95.0	390	2	D86291	hypothetical prote	621	19	95.0	431	2	C69087	dihydrolipoamide d
549	19	95.0	391	2	C70972	probable trehalose	622	19	95.0	431	2	G96987	enolase [imported]
550	19	95.0	392	2	B64017	probable RNA methy	623	19	95.0	431	2	T21594	hypothetical prote
551	19	95.0	392	2	S22580	telomere-binding p	624	19	95.0	431	2	D70185	3-hydroxy-3-methyl
552	19	95.0	393	2	E95261	serine proteinase	625	19	95.0	431	2	S37775	filamin, muscle -
553	19	95.0	393	2	C89561	hypothetical prote	626	19	95.0	432	1	S15203	glycine hydroxymet
554	19	95.0	394	2	B70411	fimbrial assembly	627	19	95.0	432	2	S08277	cyclin A - human
555	19	95.0	394	2	C81333	probable efflux pu	628	19	95.0	432	2	A02403	ATP-binding protei
556	19	95.0	394	2	S76331	hypothetical prote	629	19	95.0	433	1	B70677	cytochrome P450 Rv
557	19	95.0	395	2	S38812	cyclin A - chicken	630	19	95.0	433	2	A70465	probable GTP bindi
558	19	95.0	395	2	T20724	hypothetical prote	631	19	95.0	434	2	S30334	glycine hydroxymet
559	19	95.0	397	2	H84225	acyl-CoA dehydroge	632	19	95.0	434	2	D71480	probable flagellum
560	19	95.0	397	2	B98127	serine proteinase	633	19	95.0	436	2	G69466	3-hydroxy-3-methyl
561	19	95.0	397	2	A72315	hypothetical prote	634	19	95.0	437	1	A31752	transcription fact
562	19	95.0	398	2	C95201	hydroxymethylgluta	635	19	95.0	437	2	C64113	tetrahydrofolylpol
563	19	95.0	398	2	A98068	hypothetical prote	636	19	95.0	437	2	JC4988	high-affinity gluc
564	19	95.0	399	2	A84572	nicotinate phospho	637	19	95.0	437	2	S42111	transcription fact
565	19	95.0	399	2	H87342	glycosyl transfera	638	19	95.0	438	2	E95383	probable aminotran
566	19	95.0	399	2	F83796	multidrug-efflux t	639	19	95.0	438	2	A98161	hypothetical prote
567	19	95.0	399	2	AG2221	hypothetical prote	640	19	95.0	438	2	H86006	hypothetical prote
568	19	95.0	400	2	T03460	probable leucine/i	641	19	95.0	438	2	AB0997	high-affinity gluc
569	19	95.0	402	1	A70707	cytochrome P450 Rv	642	19	95.0	439	2	AC0462	glycerol-3-phospha
570	19	95.0	402	2	H82813	cytochrome P450-li	643	19	95.0	439	2	AC2948	hypothetical prote
571	19	95.0	403	2	H82440	ABC transporter, A	644	19	95.0	439	2	G93334	glycerol-3-phospha
572	19	95.0	403	2	G84222	NADH dehydrogenase	645	19	95.0	440	2	AC2985	conserved hypothet
573	19	95.0	405	2	A83084	probable metallope	646	19	95.0	440	2	D98298	probable sugar iso
574	19	95.0	406	2	S24788	cyclin A - bovine	647	19	95.0	441	2	C87233	probable cell inva
575	19	95.0	406	2	F84152	serine proteinase	648	19	95.0	442	2	T44655	O-acetylhomoserine
576	19	95.0	406	2	G02022	tryptophan oxygena	649	19	95.0	443	2	A99657	hypothetical prote
577	19	95.0	406	2	T30748	hypothetical prote	650	19	95.0	443	2	A85508	hypothetical prote
578	19	95.0	407	2	B71679	hypothetical prote	651	19	95.0	443	2	T32088	hypothetical prote
579	19	95.0	408	2	T08069	protein kinase, 48	652	19	95.0	444	2	C70444	biotin carboxylase
580	19	95.0	408	2	B97741	hypothetical prote	653	19	95.0	445	2	T16025	hypothetical prote
581	19	95.0	409	2	AD0492	multidrug transloc	654	19	95.0	445	2	S32036	finger protein XFG
582	19	95.0	409	2	T02776	Y4W protein - Rhi	655	19	95.0	445	2	C82140	undulin 2 - human
583	19	95.0	409	2	F86151	F22M8.10 protein -	656	19	95.0	446	2	C82140	C4-dicarboxylate t
584	19	95.0	410	2	B48585	transcription fact	657	19	95.0	447	2	C96497	glyceraldehyde-3-p
585	19	95.0	410	2	A48585	transcription fact	658	19	95.0	448	2	AF0437	conserved hypothet
586	19	95.0	411	2	H83355	hypothetical prote	659	19	95.0	449	2	D70726	probable gabt - My
587	19	95.0	411	2	AB0199	peptidase T (EC 3.	660	19	95.0	449	2	C70846	probable metC prot
588	19	95.0	412	1	RNECTA	rRNA adenyllytrans	661	19	95.0	449	2	AF2646	chemotaxis motD pr
589	19	95.0	412	2	S51760	ferredoxin reducta	662	19	95.0	449	2	E97428	chemotaxis motD pr
590	19	95.0	412	2	C91121	tRNA nucleotidyl t	663	19	95.0	450	2	T39433	glucoamylase precu
591	19	95.0	412	2	B85966	tRNA nucleotidyl t	664	19	95.0	450	2	F82402	permease VCA0904 l
592	19	95.0	412	2	T21419	hypothetical prote	665	19	95.0	450	2	T21593	hypothetical prote
593	19	95.0	413	2	C69160	phosphoenolpyruvat	666	19	95.0	450	2	C84466	hypothetical prote
594	19	95.0	413	2	T48364	transcription fact	667	19	95.0	453	2	S37935	conserved hypothet
595	19	95.0	415	2	A81920	glutamyl-tRNA redu	668	19	95.0	453	2	C83008	probable purine-bi
596	19	95.0	415	2	C81183	glutamyl-tRNA redu	669	19	95.0	455	2	B71335	hypothetical prote
597	19	95.0	415	2	I51637	cyclin A2 - Africa	670	19	95.0	456	2	T46986	probable chlorohyd
598	19	95.0	416	2	D95321	diaminopimelate de	671	19	95.0	456	2	AH0240	DNA-directed DNA p
599	19	95.0	416	2	F98095	diaminopimelate de	672	19	95.0	458	2	S75328	repSA protein - St
600	19	95.0	418	2	H72203	hypothetical prote	673	19	95.0	458	2	A12218	hypothetical prote
601	19	95.0	418	2	T35993	probable aminotran	674	19	95.0	460	2	A12218	argininosuccinate
602	19	95.0	419	2	B85035	hypothetical prote	675	19	95.0	461	1	WZTRRS	argininosuccinate
603	19	95.0	420	2	S65084	finger protein XFG	676	19	95.0	461	2	E71672	malonate hydratase
604	19	95.0	421	2	A90003	hypothetical prote	677	19	95.0	462	2	A33313	gene XGF 5.1C prot
605	19	95.0	421	2	T47393	hypothetical prote	678	19	95.0	462	2	D97826	malonate hydratase
606	19	95.0	422	2	C83063	glutamyl-tRNA redu	679	19	95.0	463	1	WZTRRS	argininosuccinate
607	19	95.0	422	2	S38501	cyclin A2 - mouse	680	19	95.0	464	1	WZTRRS	argininosuccinate
608	19	95.0	422	2	S37280	cyclin A - mouse	681	19	95.0	465	2	F70837	probable sulfatase
609	19	95.0	423	2	T19145	hypothetical prote	682	19	95.0	466	2	T36212	replication initia
610	19	95.0	424	2	I39506	citrate (si)-synth	683	19	95.0	467	2	A12231	nicotinamide nucle
611	19	95.0	424	2	H87520	hypothetical prote	684	19	95.0	467	2	G82697	hypothetical prote
612	19	95.0	426	2	H75407	transcription term	685	19	95.0	468	2	S39832	probable phosphor
613	19	95.0	427	2	AB3025	methionine gamma-1	686	19	95.0	469	1	KCPGI	interstitial colla

587	19	95.0	469	2	S74825	probable Rieske ir	760	19	95.0	513	2	A85697	Na+/H+ antiporter,
588	19	95.0	470	2	S50083	photolyase - short	761	19	95.0	513	2	G64864	Na+/H+-exchanging
589	19	95.0	471	2	S27676	mannose-1-phosphat	762	19	95.0	514	2	S46733	hypothetical prote
590	19	95.0	472	2	C82917	preprotein translo	763	19	95.0	514	2	T10559	hypothetical prote
591	19	95.0	473	2	T04738	hypothetical prote	764	19	95.0	514	2	AG0723	regulator of intra
592	19	95.0	474	2	E47677	guanosine diphosph	765	19	95.0	517	2	T29852	hypothetical prote
593	19	95.0	475	2	S22621	Phosphomannomutase	766	19	95.0	517	2	T27101	hypothetical prote
594	19	95.0	476	2	S24953	conserved hypothet	767	19	95.0	518	1	A27705	alpha-amylase (EC
595	19	95.0	477	2	T84157	hypothetical prote	768	19	95.0	521	2	T46687	p-cresol methylhyd
596	19	95.0	478	2	S75530	hydrogenase large	769	19	95.0	522	2	A31556	glucose transport
597	19	95.0	479	2	AG3210	proline dipeptidas	770	19	95.0	522	2	B64067	Na+/H+-exchanging
598	19	95.0	480	2	B49681	long-chain-fatty-a	771	19	95.0	524	2	S06920	glucose transport
599	19	95.0	481	2	G87292	hypothetical prote	772	19	95.0	524	1	T40127	IMP dehydrogenase
600	19	95.0	482	2	C64119	starch synthase (E	773	19	95.0	524	2	A31318	glucose transport
601	19	95.0	483	2	C96577	hypothetical prote	774	19	95.0	524	2	T06134	hypothetical prote
602	19	95.0	484	2	S71360	noea protein - Rhi	775	19	95.0	524	2	T09331	hypothetical prote
603	19	95.0	485	2	H95332	Noea host specific	776	19	95.0	526	2	D95205	hypothetical prote
604	19	95.0	486	2	JC4386	adenyl cyclase-a	777	19	95.0	528	2	T31459	proteochlorophyllid
605	19	95.0	487	2	T38409	adenyl cyclase-a	778	19	95.0	528	2	JC4814	Na+/H+-exchanging
606	19	95.0	488	2	T23792	hypothetical prote	779	19	95.0	528	2	AF3059	cytochrome d oxida
607	19	95.0	489	2	S15312	rfbM protein - Sal	780	19	95.0	529	2	B98227	cytochrome d oxida
608	19	95.0	490	2	AH0765	mannose-1-phosphat	781	19	95.0	532	2	T07903	tubulin delta chai
609	19	95.0	491	2	B86367	protein F26P24.16	782	19	95.0	533	2	S52046	deoxyribodipyrimid
610	19	95.0	492	2	H96729	probable alanine a	783	19	95.0	535	2	C82433	methyl-accepting c
611	19	95.0	493	2	A70190	hypothetical prote	784	19	95.0	537	2	T09755	4-coumarate-CoA li
612	19	95.0	494	2	T10095	nifs protein - Met	785	19	95.0	537	2	T09710	4-coumarate-CoA li
613	19	95.0	495	2	T47422	cellulase-like pro	786	19	95.0	537	2	AF0575	2,3-dihydroxybenzo
614	19	95.0	496	2	D82446	probable formate t	787	19	95.0	537	2	H87502	conserved hypothet
615	19	95.0	497	2	T76302	probable 2-dehydro	788	19	95.0	538	2	AD3281	IMP cyclohydrolyase
616	19	95.0	498	2	R86416	unknown protein, 3	789	19	95.0	538	2	AF2923	bifunctional purin
617	19	95.0	499	2	P95238	PTS system, membra	790	19	95.0	540	2	S76584	hypothetical prote
618	19	95.0	500	2	H83131	5-carboxy-2-hydrox	791	19	95.0	543	2	H82282	vibriobactin-speci
619	19	95.0	501	2	C96018	probable starch sy	792	19	95.0	545	2	F83280	probable chemotaxi
620	19	95.0	502	2	T49017	hypothetical prote	793	19	95.0	548	2	AB0591	fumarate hydratase
621	19	95.0	503	2	T39456	zinc finger protein	794	19	95.0	548	2	A81854	glucose-6-phosphat
622	19	95.0	504	2	A95926	probable argininos	795	19	95.0	548	2	C81089	glucose-6-phosphat
623	19	95.0	505	2	T09734	1-aminocyclopropan	796	19	95.0	548	2	T23270	acetyl choline rec
624	19	95.0	506	2	T23055	hypothetical prote	797	19	95.0	549	2	A90253	hypothetical prote
625	19	95.0	507	2	R90423	sugar transport pr	798	19	95.0	549	2	JC4508	hypothetical prote
626	19	95.0	508	2	T44576	L-2,4-diaminobuty	799	19	95.0	549	2	T24083	chaperonin - Caeno
627	19	95.0	509	2	E71486	probable s/t prote	800	19	95.0	550	2	B82330	glucose-6-phosphat
628	19	95.0	510	2	S37053	glucose-6-phosphat	801	19	95.0	550	2	T24019	hypothetical prote
629	19	95.0	511	2	T11894	NADH2 dehydrogenas	802	19	95.0	550	2	T22557	hypothetical prote
630	19	95.0	512	2	B64642	cell division prot	803	19	95.0	551	2	E84106	hypothetical prote
631	19	95.0	513	2	T01235	hypothetical prote	804	19	95.0	553	2	S03828	regulatory protein
632	19	95.0	514	2	C71873	septum formation p	805	19	95.0	553	2	AG3604	cytochrome d ubiq
633	19	95.0	515	2	H70573	hypothetical prote	806	19	95.0	555	1	SYHOMA	malate synthase (E
634	19	95.0	516	2	S10134	plasmid recombinat	807	19	95.0	556	2	C87472	steroid monooxygen
635	19	95.0	517	2	A33952	58K mobilization p	808	19	95.0	556	2	T16790	hypothetical prote
636	19	95.0	518	2	T27693	hypothetical prote	809	19	95.0	558	2	T29503	serine C-palmitoyl
637	19	95.0	519	2	AD2375	hypothetical prote	810	19	95.0	558	2	A75216	hypothetical prote
638	19	95.0	520	2	S27785	acetyl-CoA acetyl	811	19	95.0	558	2	T28881	hypothetical prote
639	19	95.0	521	2	AG3134	N-ethylameline ch	812	19	95.0	559	2	AB8813	xyran 1,4-beta-xy
640	19	95.0	522	2	T14236	NADH2 dehydrogenas	813	19	95.0	559	2	AB0984	probable membrane
641	19	95.0	523	2	F88359	protein V51H1A.5 {	814	19	95.0	559	2	H82532	conserved hypothet
642	19	95.0	524	2	R98153	N-ethylameline ch	815	19	95.0	561	2	G81842	30S ribosomal prot
643	19	95.0	525	2	B64069	fructose phosphotr	816	19	95.0	561	2	G81099	30S ribosomal prot
644	19	95.0	526	2	R86148	hypothetical prote	817	19	95.0	561	2	H69086	ABC transporter -
645	19	95.0	527	2	C91251	probable tail shea	818	19	95.0	562	2	F97697	hypothetical prote
646	19	95.0	528	2	B81823	exopolysphatase	819	19	95.0	562	2	B82351	hypothetical prote
647	19	95.0	529	2	F81080	exopolysphatase	820	19	95.0	564	2	B72660	probable type III D
648	19	95.0	530	2	G81917	probable ubiquinol	821	19	95.0	567	2	S69779	adenin A65-2 prote
649	19	95.0	531	2	B81184	ubiquinone biosynt	822	19	95.0	567	2	T16105	hypothetical prote
650	19	95.0	532	2	A82193	Sun/nucleolar prot	823	19	95.0	567	2	D72359	conserved hypothet
651	19	95.0	533	2	S27113	malonyl-CoA decarb	824	19	95.0	570	2	C90485	beta-glucuronidase
652	19	95.0	534	2	A84244	probable signaling	825	19	95.0	571	2	T06737	hypothetical prote
653	19	95.0	535	2	T16903	hypothetical prote	826	19	95.0	576	1	S33853	hydrogenase (EC 1.
654	19	95.0	536	2	C70468	phosphoribosylamin	827	19	95.0	576	2	B71959	proline-tRNA ligas
655	19	95.0	537	2	T31831	hypothetical prote	828	19	95.0	577	2	C63059	acetylacate synth
656	19	95.0	538	2	AD0475	probable magnesium	829	19	95.0	577	2	F84549	proline-tRNA ligas
657	19	95.0	539	2	G98102	hypothetical prote	830	19	95.0	580	2	T30583	probable peptidase
658	19	95.0	540	2	T14638	cytochrome P450 Cy	831	19	95.0	580	2	C86530	31 ribosomal prote
659	19	95.0	541	2	A99839	Na+/H+ antiporter	832	19	95.0	580	2	D72093	ribosomal protein

833	19	95.0	582	2	T16104	hypothetical prote	906	19	95.0	639	2	A56126	peroxisomal target
834	19	95.0	583	2	G96481	Mutator-like trans	907	19	95.0	640	1	A55073	transforming prote
835	19	95.0	583	2	T46177	villin 3 homolog T	908	19	95.0	641	2	F75553	hypothetical prote
836	19	95.0	587	2	E82431	methyl-accepting c	909	19	95.0	640	2	AD3502	dnak protein (impo
837	19	95.0	588	2	T25248	hypothetical prote	910	19	95.0	642	2	S55521	beta-fructofuranos
838	19	95.0	588	2	A43740	DG42 protein - Afr	911	19	95.0	644	2	JQ0160	3-isopropylmalate
839	19	95.0	588	2	A33460	hypothetical cytos	912	19	95.0	644	2	A45635	dnak-type molecula
840	19	95.0	589	2	A46459	macrophage-activat	913	19	95.0	645	2	JC5642	hypothetical prote
841	19	95.0	589	2	S74668	hypothetical prote	914	19	95.0	645	2	T25824	hypothetical prote
842	19	95.0	590	2	B81104	nitrate/nitrite se	915	19	95.0	646	2	B86295	transcription anti
843	19	95.0	590	2	C81911	nitrate/nitrite se	916	19	95.0	647	2	AF1488	probable thiamin p
844	19	95.0	591	2	S43506	hypothetical prote	917	19	95.0	648	2	AG0314	drebrin E (clone g
845	19	95.0	591	2	S70524	guanine nucleotide	918	19	95.0	649	2	UN0809	hypothetical prote
846	19	95.0	592	2	S64923	probable membrane	919	19	95.0	650	2	F72540	beta-fructofuranos
847	19	95.0	592	2	T45827	pectinesterase-lik	920	19	95.0	651	2	T12083	dnak-type molecula
848	19	95.0	598	2	B71095	hypothetical prote	921	19	95.0	651	2	JC4610	hypothetical prote
849	19	95.0	598	2	T32430	hypothetical prote	922	19	95.0	652	2	T14763	paraaporal crystal
850	19	95.0	602	2	D89331	probable DNA topoi	923	19	95.0	652	2	I39811	transferrin-bindin
851	19	95.0	602	2	T01360	hypothetical prote	924	19	95.0	654	2	S70905	peptidase (impor
852	19	95.0	603	2	T39674	hypothetical dnaj	925	19	95.0	654	2	AD3183	hypothetical prote
853	19	95.0	604	2	F87936	protein M01G12.12	926	19	95.0	656	1	E70107	transferrin-bindin
854	19	95.0	604	2	T23669	hypothetical prote	927	19	95.0	660	2	S70904	ATP-dependent DNA
855	19	95.0	605	2	I39837	dnak-type molecula	928	19	95.0	669	2	C82984	hypothetical prote
856	19	95.0	605	2	D84687	hypothetical prote	929	19	95.0	669	2	C96630	hypothetical prote
857	19	95.0	606	2	S46833	transfrase - Stre	930	19	95.0	670	2	D86457	hypothetical prote
858	19	95.0	607	2	T35710	dnak protein (impo	931	19	95.0	670	2	AH2425	probable chemotaxi
859	19	95.0	607	2	S86744	dnak-type molecula	932	19	95.0	672	2	S61157	bacterio-opsin act
860	19	95.0	607	2	S39342	dnak protein (impo	933	19	95.0	674	2	E84300	hypothetical prote
861	19	95.0	607	2	B95060	hypothetical prote	934	19	95.0	674	2	T22733	NADPH-ferrinemopro
862	19	95.0	607	2	G37928	hypothetical prote	935	19	95.0	678	1	RDRT04	growth arrest-spec
863	19	95.0	607	2	A58898	probable sensor ki	936	19	95.0	680	2	B48089	peptidyl-dipectida
864	19	95.0	610	2	S41335	hypothetical prote	937	19	95.0	680	2	A42297	dipectidyl carboxy
865	19	95.0	610	2	C89939	conserved hypoteth	938	19	95.0	680	2	AH0678	hypothetical prote
866	19	95.0	610	2	G69130	probable ABC-type	939	19	95.0	684	2	T25602	hypothetical prote
867	19	95.0	612	2	A70756	dnak-type molecula	940	19	95.0	685	2	C64985	hypothetical prote
868	19	95.0	613	2	T43738	class I heat-shock	941	19	95.0	685	2	C64428	hypothetical prote
869	19	95.0	613	2	AE1621	class I heat-shock	942	19	95.0	688	2	B64409	hypothetical prote
870	19	95.0	613	2	A11258	protein kinase pkn	943	19	95.0	688	2	B64409	hypothetical prote
871	19	95.0	619	2	S77221	hypothetical prote	944	19	95.0	688	2	B64409	hypothetical prote
872	19	95.0	619	2	A84411	hypothetical prote	945	19	95.0	690	2	T11749	transferrin - Atla
873	19	95.0	621	2	T06717	phosphothricin-t	946	19	95.0	691	2	T32748	conserved hypoteth
874	19	95.0	622	2	PT0059	two-component sens	947	19	95.0	691	2	A22766	hypothetical prote
875	19	95.0	622	2	E83160	paraaporal crystal	948	19	95.0	691	2	H97546	transferrin precu
876	19	95.0	622	2	S17402	hypothetical prote	949	19	95.0	694	1	TRFBP	transferrin - pig
877	19	95.0	624	2	C83389	transferrin-bindin	950	19	95.0	696	1	S01384	hypothetical prote
878	19	95.0	625	2	D64107	asparagine synthas	951	19	95.0	696	1	E36831	hypothetical prote
879	19	95.0	625	2	D86903	probable periplasm	952	19	95.0	696	2	H83024	hypothetical prote
880	19	95.0	625	2	G96976	carnitine O-acetyl	953	19	95.0	698	1	TRHUP	asmA protein VC103
881	19	95.0	627	2	S53369	hypothetical prote	954	19	95.0	703	2	G32249	carbonic anhydrase
882	19	95.0	627	2	S69958	carnitine O-acetyl	955	19	95.0	704	2	I47228	phenylalanine ammo
883	19	95.0	627	2	D71974	hypothetical prote	956	19	95.0	705	2	JC5873	hypothetical prote
884	19	95.0	628	1	H64390	carbon-monoxide de	957	19	95.0	706	2	S33761	transferrin precu
885	19	95.0	628	2	B64534	hypothetical prote	958	19	95.0	707	2	S29029	phenylalanine ammo
886	19	95.0	631	1	NPV217	nucleoside-triphos	959	19	95.0	707	2	S60598	phenylalanine ammo
887	19	95.0	631	2	H36847	nucleoside-triphos	960	19	95.0	708	1	TVVPT4	large T antigen -
888	19	95.0	631	2	T28539	hypothetical prote	961	19	95.0	708	2	T14255	phenylalanine ammo
889	19	95.0	631	2	C72163	Oil protein - vari	962	19	95.0	708	2	T10909	conserved hypoteth
890	19	95.0	631	2	T37384	nucleoside triphos	963	19	95.0	710	2	JC5872	phenylalanine ammo
891	19	95.0	631	2	S70908	transferrin-bindin	964	19	95.0	712	2	E33226	phenylalanine ammo
892	19	95.0	631	2	T30702	hypothetical prote	965	19	95.0	712	2	T03663	hercylisin secretio
893	19	95.0	634	2	S32349	probable SNF2-type	966	19	95.0	712	2	T01858	hypothetical prote
894	19	95.0	635	2	B64112	dnak-type molecula	967	19	95.0	714	2	A41464	fatty-acid beta-ox
895	19	95.0	635	2	B82273	dnak protein VC085	968	19	95.0	715	1	JX0199	fatty-acid oxidati
896	19	95.0	635	2	AF1881	1-Deoxy-xylulose 5	969	19	95.0	715	1	G83269	transforming prote
897	19	95.0	636	1	TVMSMB	transforming prote	970	19	95.0	715	4	TVMSMY	glycyl-tRNA synthe
898	19	95.0	637	2	A47042	dnak-type molecula	971	19	95.0	716	2	AH2319	conserved hypoteth
899	19	95.0	637	2	B83052	DNA repair helicase	972	19	95.0	716	2	D98855	hypothetical prote
900	19	95.0	637	2	E75044	dnak-type molecula	973	19	95.0	716	2	G82627	probable glgX prot
901	19	95.0	638	1	IQECDK	heat shock protein	974	19	95.0	721	2	A70764	hypothetical prote
902	19	95.0	638	2	F90630	dnak-type molecula	975	19	95.0	722	2	S76367	hypothetical prote
903	19	95.0	638	2	F85481	dnak-type molecula	976	19	95.0				
904	19	95.0	638	2	AE0503	Dnak protein (heat	977	19	95.0				
905	19	95.0	639	2	S72163	methyl-accepting c	978	19	95.0				

979 19 95.0 724 2 B85045 probable calcium c  
980 19 95.0 725 1 Z2BPA4 rapid lysis protei  
981 19 95.0 725 2 S52990 phenylalanine ammo  
982 19 95.0 725 2 G84787 phenylalanine ammo  
983 19 95.0 725 2 A10147 probable ATP-depen  
984 19 95.0 725 2 A91038 hypothetical prote  
985 19 95.0 729 2 A90808 probable membrane  
986 19 95.0 738 2 A87516 dipeptidyl peptida  
987 19 95.0 742 2 D85882 hypothetical prote  
988 19 95.0 744 2 T13048 NADH2 dehydrogenas  
989 19 95.0 746 2 S31816 genome polyprotein  
990 19 95.0 748 2 G96458 hypothetical prote  
991 19 95.0 754 2 T25551 hypothetical prote  
992 19 95.0 756 2 T04187 subtilisin-like pr  
993 19 95.0 761 1 TVHUMB transforming prote  
994 19 95.0 763 2 S23457 polysulfide reduct  
995 19 95.0 764 2 S49849 aconitate hydratase  
996 19 95.0 768 2 H82367 sensory box/GGDEF  
997 19 95.0 768 2 H65013 yfeA protein - Esc  
998 19 95.0 772 2 T05838 subtilisin-like pr  
999 19 95.0 774 2 G71308 probable aminopept  
1000 19 95.0 781 2 A69312 DNA polymerase B1

## ALIGNMENTS

RESULT 1  
S39049  
cytotoxin-binding protein - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 18-Feb-1994 #sequence\_revision 19-Apr-1996 #text\_change 16-Feb-1997  
C:Accession: S39049  
R:Rutiz, F.; Mohr, M.; Grimmig, M.; Leidl, R.; Linder, D.  
Eur. J. Biochem. 217, 1123-1128, 1993  
A>Title: Pseudomonas aeruginosa cytotoxin-binding protein in rabbit erythrocyte membrane  
A:Reference number: S39049; MUID:94039134; PMID:7693466  
A:Accession: S39049  
A>Status: Preliminary  
A:Molecule type: protein  
A:Residues: 1-20 <LUT>  
C:Superfamily: lens fiber membrane major intrinsic protein

Query Match 95.0%; Score 19; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 14 VAEF 17

RESULT 2  
G64625  
hypothetical protein HP0847 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: G64625  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: G64625  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-33 <TOM>  
A:Cross-references: GB:AE000596; GB:AE000511; NID:g2313982; PIDN:AAD07920.1; PID:g231400

Query Match 95.0%; Score 19; DB 2; Length 33;  
Best Local Similarity 100.0%; Pred. No. 88;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 15 VAEF 18

RESULT 3  
E64618  
hypothetical protein HP0789 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: E64618  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: E64618  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-48 <TOM>  
A:Cross-references: GB:AE000591; GB:AE000511; NID:g2313918; PIDN:AAD07847.1; PID:g231392

Query Match 95.0%; Score 19; DB 2; Length 48;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 15 VAEF 18

RESULT 4  
A42375  
hypothetical protein (lon 5' region) - Bacillus brevis  
C:Species: Bacillus brevis  
C>Date: 10-Jul-1992 #sequence\_revision 10-Jul-1992 #text\_change 15-Oct-1999  
C:Accession: A42375; I39873  
R:Ito, K.; Uda, S.; Yamagata, H.  
J. Bacteriol. 174, 2281-2287, 1992  
A>Title: Cloning, characterization, and inactivation of the Bacillus brevis lon gene.  
A:Reference number: A42375; MUID:92202157; PMID:1551846  
A:Accession: A42375  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-52 <ITO>  
A:Cross-references: GB:D00863; NID:g216293; PIDN:BA00736.1; PID:g303507

Query Match 95.0%; Score 19; DB 2; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 47 VAEF 50

RESULT 5  
E85745  
unknown protein encoded within prophage CP-933R [imported] - Escherichia coli (strain O1  
C:Species: Escherichia coli  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: E85745  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: E85745  
A>Status: preliminary

```

Molecule type: DNA
;Residues: 1-54 <STO>
;Cross-references: GB:AE005174; NID:gl2515374; PIDN:RAG56425.1; GSPDB:GN00145; UWGP:223
;Experimental source: strain O157:H7, substrain EDL933
;Genetics:
;Gene: Z2382

Query Match      95.0%; Score 19; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y      2 VAEF 5
      |||||
b     32 VAEF 35

RESULT 6
69386
observed hypothetical protein AF1090 - Archaeoglobus fulgidus
;Species: Archaeoglobus fulgidus
;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 11-Jan-2000
;Accession: A69386
;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
;Glodek, A.; Zhou, L.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
ature 390, 364-370, 1997
;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
mith, H.O.; Woese, C.R.; Venter, J.C.
;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
;Reference number: A69250; MUID:98049343; PMID:9389475
;Accession: A69386
;Status: preliminary; nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-59 <KLE>
;Cross-references: GB:AE001028; GB:AE000782; NID:G2689351; PIDN:AAB90157.1; PID:G264950
;Superfamily: Methanococcus jannaschii hypothetical protein MJ0975

Query Match      95.0%; Score 19; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y      2 VAEF 5
      |||||
b     51 VAEF 54

RESULT 7
71349
robable preprotein translocase subunit (secE) - syphilis spirochete
;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
;Accession: E71349
;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
son, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uterback, T.; McDo
hey, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
cience 281, 375-388, 1998
;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
;Reference number: A71250; MUID:9832770; PMID:9665876
;Accession: E71349
;Status: preliminary; nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-59 <COL>
;Cross-references: GB:AE001205; GB:AE000520; NID:G3322501; PIDN:AAC65223.1; PID:G332250
;Experimental source: strain Nichols
;Genetics:
;Gene: TP0235

Query Match      95.0%; Score 19; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y      2 VAEF 5
      |||||

```

---

```

Molecule type: DNA
;Residues: 1-63 <ARI>
;Cross-references: DBJ:D30759; NID:G551628; PIDN:BA06419.1; PID:d1006989; PID:ig551629
;Note: the source was designated as Streptomyces thermotolerans
;Genetics:
;Gene: carA
;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
;Keywords: ATP
F:1-55/Domain: ATP-binding cassette homology (fragment) <ABC2>

Query Match      95.0%; Score 19; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY     2 VAEF 5
      |||||
Db     57 VAEF 60

RESULT 8
JC4002
carboxymycin resistance protein carA - Streptomyces sp. (fragment)
;Species: Streptomyces sp.
;Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 20-Sep-1999
;Accession: JC4002
R;Arisawa, A.; Tsunekawa, H.; Okamura, K.; Okamoto, R.
Biosci. Biotechnol. Biochem. 59, 582-588, 1995
;Title: Nucleotide sequence analysis of the carboxymycin biosynthetic genes including the
A;Reference number: JC4001; MUID:95290751; PMID:7772821
A;Accession: JC4002
A;Molecule type: DNA
A;Residues: 1-63 <ARI>
A;Cross-references: DBJ:D30759; NID:G551628; PIDN:BA06419.1; PID:d1006989; PID:ig551629
A;Note: the source was designated as Streptomyces thermotolerans
;Genetics:
;Gene: carA
;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
;Keywords: ATP
F:1-55/Domain: ATP-binding cassette homology (fragment) <ABC2>

Query Match      95.0%; Score 19; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY     2 VAEF 5
      |||||
Db     57 VAEF 60

RESULT 9
G64007
hypothetical protein HI0451 - Haemophilus influenzae (strain Rd KW20)
;Species: Haemophilus influenzae
;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
;Accession: G64007
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
;Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
;D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: G64007
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-63 <TIGR>
A;Cross-references: GB:U32728; GB:L42023; NID:G1573425; PIDN:AAC22109.1; PID:G1573437; T

Query Match      95.0%; Score 19; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY     2 VAEF 5
      |||||
Db     27 VAEF 30

RESULT 10
IS1350
transferrin - Atlantic salmon (fragment)
;Species: Salmo salar (Atlantic salmon)
;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
;Accession: IS1350
R;Kvingedal, A.M.
Gene 150, 335-339, 1994
A;Title: Characterization of the 5' region of the Atlantic salmon (Salmo salar) transfer
A;Reference number: IS1350; MUID:95121925; PMID:7821802
A;Accession: IS1350
A;Status: preliminary; translated from GB/EMBL/DDBJ

```

A:Molecule type: DNA  
 A:Residues: 1-64 <KVI>  
 A:Cross-references: GB:L26903; NID:g598395; PIDN:AAC42221.1; PID:g598396  
 C:Genetics:  
 A:Gene: If  
 A:Introns: 14/1  
 C:Superfamily: transferrin; transferrin repeat homology

Query Match 95.0%; Score 19; DB 2; Length 64;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 DB 45 VAEF 48

## RESULT 11

D97045  
 Hypothetical protein CAC1179 [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum  
 C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 C:Accession: D97045

R: Noll, J.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: D97045

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-74 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK79151.1; PID:g15024100; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC624

C:Genetics:

A:Gene: CAC1179

Query Match 95.0%; Score 19; DB 2; Length 74;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 DB 5 VAEF 8

## RESULT 12

S07277  
 Gene D protein - phage 186 (fragment)

C:Species: phage 186

C:Date: 28-May-1993 #sequence\_revision 28-May-1993 #text\_change 17-Mar-1999

C:Accession: S07277; S25273

R: Kallionis, B.; Dodd, I.B.; Egan, J.B.

J. Mol. Biol. 191, 199-203, 1986

A:Title: Control of gene expression in the P2-related template coliphages. III. DNA sequence

A:Reference number: S07277; MUID:87112711; PMID:3806670

A:Accession: S07277

A:Molecule type: DNA

A:Residues: 1-75 <KAL>

A:Cross-references: EMBL:X04449; NID:g15549; PID:g15550

R: Bibb, J.A.; Gregory, S.L.; Egan, J.B.

Mol. Microbiol. 6, 2643-2650, 1992

A:Title: Control of gene expression in the temperate coliphage 186. X. The cI repressor

A:Reference number: S25273; MUID:93078618; PMID:1447973

A:Accession: S25273

A:Molecule type: DNA

A:Residues: 1-75 <DTB>

A:Cross-references: EMBL:X04449; NID:g15549; PID:g15550

Query Match 95.0%; Score 19; DB 2; Length 75;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 DB 5 VAEF 8

## RESULT 13

Ti8034

Hypothetical protein A532L - Chlorella virus PBCV-1

C:Species: Chlorella virus PBCV-1

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: Ti8034

R: Graves, M.V.; Van Etten, J.L.

submitted to the EMBL Data Library, May 1999

A:Reference number: Z18806

A:Accession: Ti8034

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-79 <GRA>

A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96899.1

A:Experimental source: specific host Chlorella strain NC64A

C:Genetics:

A:Note: A532L

Query Match 95.0%; Score 19; DB 2; Length 79;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 DB 49 VAEF 52

## RESULT 14

T27605

Hypothetical protein ZC477.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T27605

R: Du, Z.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid ZC477.

A:Reference number: Z20392

A:Accession: T27605

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-79 <DUZ>

A:Cross-references: EMBL:U40802; PIDN:AAA81506.1; CESP:ZC477.6

C:Genetics:

A:Gene: CESP:ZC477.6

A:Introns: 51/2

Query Match 95.0%; Score 19; DB 2; Length 79;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 DB 74 VAEF 77

## RESULT 15

F83922

Hypothetical protein BH2182 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C:Accession: F83922

R: Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: F83922

A>Status: preliminary

A:Molecule type: DNA



```

;Residues: 1-84 <STO>
;Cross-references: GB:AP001514; GB:BA000004; NID:G10174613; PIDN:BA05901.1; GSPDB:GN00
;Experimental source: strain C-125
;Genetics:
;Gene: BH2182

Query Match          95.0%; Score 19; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/      2 VAEF 5
      ||||
      38 VAEF 41

38725
;Title: transferrin - sheep (fragments)
;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
;Date: 20-Sep-1991 #sequence_revision 20-Sep-1991 #text_change 05-Dec-1997
;Accession: A38725
;Author: Tu, G.F.; Achen, M.G.; Aldred, A.R.; Southwell, B.R.; Schreiber, G.
;Biol. Chem. 266, 6201-6209, 1991
;Title: The distribution of cerebral expression of the transferrin gene is species spec
;Reference number: A38725; MUID:91177867; PMID:1848850
;Accession: A38725
;Status: preliminary
;Molecule type: mRNA
;Residues: 1-87 <TUA>
;Cross-references: GB:M64692
;Superfamily: transferrin; transferrin repeat homology
;Keywords: duplication

Query Match          95.0%; Score 19; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/      2 VAEF 5
      ||||
      13 VAEF 16

38716
;Title: Isosomal protein S15 - Haemophilus influenzae (strain Rd KW20)
;Species: Haemophilus influenzae
;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
;Accession: H64116; E64125
;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
;Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
;D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
;Science 269, 496-512, 1995
;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
;Reference number: A64000; MUID:95350630; PMID:7542800
;Accession: H64116
;Status: nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-89 <TIG1>
;Cross-references: GB:U32812; GB:U42023; NID:G1574784; PIDN:AAC22973.1; PID:G1574788; T
;Note: named as homolog to a protein from Escherichia coli
;Accession: E64125
;Status: nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-89 <TIG2>
;Cross-references: GB:U32825; GB:U42023; NID:G1574307; PIDN:AAC23117.1; PID:G1574309; T
;Note: named by homology to a protein from Escherichia coli
;Genetics:
;Superfamily: two copies of this gene are found in the Haemophilus influenzae chromosome
;Note: family: Escherichia coli ribosomal protein S15; eubacterial ribosomal protein S15
;23-99/Domain: eubacterial ribosomal protein S15 homology <BS15>

Query Match          95.0%; Score 19; DB 1; Length 89;

```

---

```

Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
      ||||
      12 VAEF 15

Db

RESULT 18
P86818
hypothetical protein YPjB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: F86818
R:Polotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssi
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: F86818
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-89 <STO>
A:Cross-references: GB:AE005176; PID:G12724551; PIDN:AAK05648.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: YPjB

Query Match          95.0%; Score 19; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
      ||||
      68 VAEF 71

Db

RESULT 19
T15018
hypothetical protein Y1103 - Yersinia pestis plasmid pMT1
C:Species: Yersinia pestis
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15018
R:Lindler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.
Infect. Immun. 66, 5731-5742, 1998
A:Title: Complete DNA sequence and detailed analysis of the Yersinia pestis KIM5 plasmid
A:Reference number: Z18268; MUID:99043898; PMID:9826348
A:Accession: T15018
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-89 <LIN>
A:Cross-references: EMBL:AF074611; NID:G3883003; PID:G3883101; PIDN:AAC82761.1
C:Genetics:
A:Gene: Y1103
A:Genome: plasmid pMT1

Query Match          95.0%; Score 19; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
      ||||
      3 VAEF 6

Db

RESULT 20
AB3053
hypothetical protein Atu4039 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AB3053
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
;Karp, P.; Romero, P.; Zhang, S.
;Karp, P.; Romero, P.; Zhang, S.

```

Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.E.W.  
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AB3053  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-96 <KUR>  
A:Cross-references: GB:AB008689; PIDN:AAU44840.1; PID:g17742484; GSPDB:GN00187  
A:Experimental source: strain C58 (Dupont)  
A:Genetics:  
A:Gene: Atu4039  
A:Map position: linear chromosome

Query Match 95.0%; Score 19; DB 2; Length 96;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
DB 55 VAEF 58

RESULT 21  
AC1397  
conserved hypothetical protein lmo2579 [imported] - *Listeria monocytogenes* (strain EGD-e)  
C:Species: *Listeria monocytogenes*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AC1397  
R:Glaser, P.; Prangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 845-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative Genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AC1397  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-97 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAD00657.1; PID:g16412067; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
A:Genetics:  
A:Gene: lmo2579

Query Match 95.0%; Score 19; DB 2; Length 97;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
DB 15 VAEF 18

RESULT 22  
E83244  
conserved hypothetical protein PA3202 [imported] - *Pseudomonas aeruginosa* (strain PAO1)  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: E83244  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen.  
A:Reference number: AB2950; MUID:20437337; PMID:10384043  
A:Accession: E83244  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-99 <STO>  
A:Cross-references: GB:AE004744; GB:AE004091; NID:g9949317; PIDN:AAG06590.1; GSPDB:GN001

A:Experimental source: strain PAO1  
A:Genetics:  
A:Gene: PA3202

Query Match 95.0%; Score 19; DB 2; Length 99;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
DB 62 VAEF 65

RESULT 23  
H90297  
conserved hypothetical protein [imported] - *Sulfolobus solfataricus*  
C:Species: *Sulfolobus solfataricus*  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
C:Accession: H90297  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Petig, X.; Thi-Ngoc, H.P.; Redder, P.; Rett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
Submitted to GenBank, April 2001  
A:Description: *Sulfolobus solfataricus* complete genome.  
A:Reference number: A99139  
A:Accession: H90297  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-101 <KUR>  
A:Cross-references: GB:AE006641; NID:g13814625; PIDN:AAK41639.1; GSPDB:GN00155  
A:Genetics:  
A:Gene: SS01404

Query Match 95.0%; Score 19; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
DB 20 VAEF 23

RESULT 24  
TI0097  
nifX protein - *Methanococcus maripaludis*  
C:Species: *Methanococcus maripaludis*  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: TI0097  
R:Kessler, P.S.; McLarnan, J.; Leigh, J.A.  
J. Bacteriol. 179, 541-543, 1997  
A:Title: Nitrogenase phylogeny and the molybdenum dependence of nitrogen fixation in *Met*  
A:Reference number: Z16944; MUID:97144542; PMID:8990309  
A:Accession: TI0097  
A>Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-106 <KES>  
A:Cross-references: EMBL:U75887; NID:g1666882; PIDN:AAC45519.1; PID:g1666890  
A:Experimental source: strain JJ  
A:Genetics:  
A:Gene: nifX

Query Match 95.0%; Score 19; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
DB 35 VAEF 38

RESULT 25  
C64490  
hypothetical protein MJ1524 - *Methanococcus jannaschii*

```

;Species: Methanococcus jannaschii
;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 12-Jun-2003
;Accession: C64490
;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
;Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
;Son, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
;Science 273, 1058-1073, 1996
;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
;Reference number: A64300; MUID:96337999; PMID:8688087
;Accession: C64490
;Status: preliminary; nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-108 <BUL>
;Cross-references: GB:U67593; GB:L77117; NID:g2826427; PIDN:AAB99550.1; PID:g1592154; T
;Genetics:
;Map position: FOR1501406-1501732
;Superfamily: uncharacterized conserved protein MJ1524
Query Match 95.0%; Score 19; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Y 2 VAEF 5
C 52 VAEF 55

RESULT 26
70949
robable regulator - Mycobacterium tuberculosis (strain H37RV)
;Species: Mycobacterium tuberculosis
;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
;Accession: E70949
;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
;Comor, R.; Davies, R.; Delvin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
;ajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skelton, S.; Squares, S.
;ature 393, 537-544, 1998
;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
;Reference number: A70500; MUID:98295987; PMID:9634230
;Accession: E70949
;Status: preliminary; nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-109 <COL>
;Cross-references: GB:AL021646; GB:AL123456; NID:g3242278; PIDN:CAA16648.1; PID:g282759
;Experimental source: strain H37RV
;Genetics:
;Gene: Rv3183

Query Match 95.0%; Score 19; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Y 2 VAEF 5
b 97 VAEF 100

RESULT 27
17582
ypothetical protein A92L - Chlorella virus PBCV-1
;Species: Chlorella virus PBCV-1
;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
;Accession: T17582
;Graves, M.V.; Van Etten, J.L.
;Submitted to the EMBL Data Library, May 1999
;Reference number: Z18806
;Accession: T17582
;Status: preliminary; translated from GB/EMBL/DBJ
;Molecule type: DNA
;Residues: 1-111 <GRA>
;Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96460.1

;Experimental source: specific host Chlorella strain NC64
;Genetics:
;C/Gene: A92L
;C/Superfamily: Chlorella virus PBCV-1 hypothetical protein A92L
Query Match 95.0%; Score 19; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Y 2 VAEF 5
Db 17 VAEF 20

RESULT 28
B83431
type III export protein PscI PA1722 [imported] - Pseudomonas aeruginosa (strain PA01)
;Species: Pseudomonas aeruginosa
;C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
;C/Accession: B83431
;R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri
;adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
;Lory, S.; Olson, M.V.
;Nature 406, 959-964, 2000
;A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathog
;A/Reference number: A82950; MUID:20437337; PMID:10984043
;A/Accession: B83431
;A/Status: preliminary
;A/Molecule type: DNA
;A/Residues: 1-112 <STO>
;A/Cross-references: GB:AE004598; GB:AE004091; NID:g9947687; PIDN:AAG05111.1; GSPDB:GN001;
;A/Experimental source: strain PA01
;C/Genetics:
;A/Gene: pscI; PA1722
;C/Superfamily: Yersinia enterocolitica plasmid pYV vitC-region hypothetical protein yscI
Query Match 95.0%; Score 19; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Y 2 VAEF 5
Db 28 VAEF 31

RESULT 29
A55872
kedarcidin [validated] - Streptococcus sp. (ATCC 53650)
;Alternate names: kedarcidin apoprotein
;C/Species: Streptococcus sp.
;A/Variety: strain L585-6 (ATCC 53650)
;C/Date: 21-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 15-Sep-2000
;C/Accession: A55872
;R/Hofstead, S.J.; Watson, J.A.; Malacko, A.R.; Marquardt, H.
;J. Antibiot. 45, 1250-1254, 1992
;A/Title: Kedarcidin, a new chromoprotein antitumor antibiotic. II. Isolation, purificati
;A/Reference number: A55872; MUID:93015257; PMID:1399845
;A/Accession: A55872
;A/Molecule type: protein
;A/Residues: 1-114 <HOF>
;A/Note: sequence extracted from NCBI backbone (NCBI:118354)
;R/Constantine, K.L.; Colson, K.L.; Wittekind, M.; Friedrichs, M.S.; Zein, N.; Tuttle, J.;
;ler, L.
;Submitted to the Brookhaven Protein Data Bank, June 1994
;A/Reference number: A52519; PDB:1AKP
;A/Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residues:
;R/Zein, N.; Casazza, A.M.; Doyle, T.W.; Leet, J.E.; Schroeder, D.R.; Solomon, W.; Nadler,
;Proc. Natl. Acad. Sci. U.S.A. 90, 8009-8012, 1993
;A/Title: Selective proteolytic activity of the antitumor agent kedarcidin.
;A/Reference number: A58601; MUID:93376732; PMID:8367457
;A/Contents: annotation; protein activity
;A/Note: the apoprotein may have proteolytic activity
;C/Superfamily: macromonocin

```

2;Keywords: antibiotic; pigment binding  
2;37-47,88-95/Disulfide bonds: #status experimental

Query Match 95.0%; Score 19; DB 2; Length 114;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
49 VAEF 52

RESULT 30

97414

hypothetical protein AGR\_C\_804 [imported] - Agrobacterium tumefaciens (strain C58, Cere

2;Species: Agrobacterium tumefaciens

2;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002

2;Accession: D97414

2;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

2;A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.

2;Science 294, 2323-2328, 2001

2;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

2;Reference number: A97359; MUID:21608551; PMID:11743194

2;Accession: D97414

2;Status: preliminary

2;Molecule type: DNA

2;Residues: 1-114 <KUR>

2;Cross-references: GB:AE007869; PIDN:AAK86269.1; PID:g15155379; GSPDB:GN00169

2;Genetics:

2;Gene: AGR\_C\_804

2;Map position: circular chromosome

Query Match 95.0%; Score 19; DB 2; Length 114;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
40 VAEF 43

RESULT 31

975218

hypothetical protein PAB0185 - Pyrococcus abyssi (strain Orsay)

2;Species: Pyrococcus abyssi

2;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999

2;Accession: D75218

2;Anonymous, Genoscope

2;submitted to the EMBL Data Library, July 1999

2;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

2;Reference number: A75001

2;Accession: D75218

2;Status: preliminary

2;Molecule type: DNA

2;Residues: 1-115 <KAW>

2;Cross-references: GB:AJ249283; GB:AL096836; NID:g5457433; PIDN:CAB49195.1; PID:e151508

2;Experimental source: strain Orsay

2;Genetics:

2;Gene: PAB0185

Query Match 95.0%; Score 19; DB 2; Length 115;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
101 VAEF 104

RESULT 32

IPAF

insulin precursor - American goosfish

2;Species: Lophius americanus (American goosfish)

2;Date: 31-Mar-1981 #sequence\_revision 31-Mar-1981 #text\_change 16-Jul-1999

2;Accession: A01608

2;Hobart, P.M.; Shen, L.P.; Crawford, R.; Pictet, R.L.; Rutter, W.J.

2;Science 210, 1360-1363, 1980

2;Title: Comparison of the nucleic acid sequence of anglerfish and mammalian insulin mRN

2;Reference number: A94254; MUID:81056434; PMID:7001633

2;Accession: A01608

2;Molecule type: mRNA

2;Residues: 1-116 <HOB>

2;Cross-references: GB:V00634; GB:J00934; NID:g64025; PIDN:CAA23907.1; PID:g64026

2;Superfamily: insulin

2;Keywords: hormone; pancreas

2;F1-24/Domain: signal sequence #status predicted <SIG>

2;F25-54/Domain: insulin chain B #status predicted <SCH>

2;F25-54,96-116/Product: insulin #status predicted <MAT>

2;F56-93/Domain: connecting peptide #status predicted <CFEP>

2;F95-116/Domain: insulin chain A #status predicted <ACH>

2;F32-102,44-115,101-106/Disulfide bonds: #status predicted

Query Match 95.0%; Score 19; DB 1; Length 116;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
80 VAEF 83

RESULT 33

T48682

hypothetical protein DKFZp761N05121.1 - human (fragment)

2;Species: Homo sapiens (man)

2;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 05-May-2000

2;Accession: T48682

2;Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.

2;submitted to the Protein Sequence Database, April 2000

2;Reference number: Z24535

2;Accession: T48682

2;Status: preliminary

2;Molecule type: mRNA

2;Residues: 1-118 <AAA>

2;Cross-references: EMBL:AL353932

2;Experimental source: adult amygdala; clone DKFZp761N05121

2;Genetics:

2;Note: DKFZp761N05121.1

Query Match 95.0%; Score 19; DB 2; Length 118;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
15 VAEF 18

RESULT 34

F95327

hypothetical protein Sma0974 [imported] - Sinorhizobium meliloti (strain 1021) magapiasm

2;Species: Sinorhizobium meliloti

2;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

2;Accession: F95327

2;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows

2;Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.

2;Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

2;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot

2;Reference number: A95262; MUID:21396509; PMID:11481432

2;Accession: F95327

2;Status: preliminary

2;Molecule type: DNA

2;Residues: 1-119 <KUR>

2;Cross-references: GB:AE006469; PIDN:AAK65184.1; PID:g14523629; GSPDB:GN00165

2;Experimental source: strain 1021, megaplasmid pSyma

2;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

la, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
 ..; Hyman, R.W.; Jones, T.  
 :ence 293, 668-672, 2001  
 Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure,  
 bault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
 Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
 Reference number: A96039; MUID:21368234; PMID:11474104  
 Contents: annotation  
 Genetics:  
 Gene: SMA0974  
 Genome: plasmid

Query Match 95.0%; Score 19; DB 2; Length 119;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 ||||  
 26 VAEF 29

RESULT 35  
 12861  
 njugal transfer protein Xfa0006 [imported] - Xylella fastidiosa (strain 9a5c)  
 Species: Xylella fastidiosa  
 Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 Accession: F82861  
 anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen-  
 ture 406, 151-157, 2000  
 Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
 Reference number: A82515; MUID:20365717; PMID:10910347  
 Note: for a complete list of authors see reference number A59328 below  
 Accession: F82861  
 Status: preliminary  
 Molecule type: DNA  
 Residues: 1-122 <SIM>  
 Cross-references: GB:AB003851; NID:G9112238; PIDN:AAF85575.1; GSPDB:GN00130; XFSC:XFA0  
 Experimental source: strain 9a5c  
 Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A-  
 -Neco, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
 bmitted to GenBank, June 2000  
 Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
 D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
 ado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, H  
 Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
 drigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
 Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
 Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
 Reference number: A59328  
 Contents: annotation  
 Genetics:  
 Gene: Xfa0006  
 Genome: plasmid  
 Note: plasmid pXFS.1

Query Match 95.0%; Score 19; DB 2; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 ||||  
 32 VAEF 35

RESULT 36  
 59129  
 nma-carboxymuconolactone decarboxylase - Methanobacterium thermoautotrophicum (strain  
 Species: Methanobacterium thermoautotrophicum  
 Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Jun-1999  
 Accession: B69129  
 Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;  
 Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A>Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: functi  
 A:Reference number: A69000; MUID:98037514; PMID:9371463  
 A:Accession: B69129  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-125 <MTH>  
 A:Cross-references: GB:AE000810; GB:AE000666; NID:G2621277; PIDN:AAB84740.1; PID:G262128;  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH234  
 C:Superfamily: 4-carboxymuconolactone decarboxylase

Query Match 95.0%; Score 19; DB 2; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 ||||  
 37 VAEF 40

RESULT 37  
 D70730  
 hypochemical protein Rv3675 - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 C:Accession: D70790  
 R;Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
 ; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
 Rajandram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A>Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: D70790  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-125 <COL>  
 A:Cross-references: GB:AL022121; GB:AL123456; NID:G3261559; PIDN:CAA17997.1; PID:el26453;  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: Rv3675

Query Match 95.0%; Score 19; DB 2; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 ||||  
 42 VAEF 45

RESULT 38  
 B81368  
 hypochemical protein Cj0393c [imported] - Campylobacter jejuni (strain NCTC 11168)  
 C:Species: Campylobacter jejuni  
 C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
 C:Accession: B81368  
 R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
 C.W.; Quail, M.; Rajandram, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell  
 Nature 403, 665-668, 2000  
 A>Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hyp  
 A:Reference number: A81250; MUID:20150912; PMID:10688204  
 A:Accession: B81368  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-126 <PAR>  
 A:Cross-references: GB:AL139076; GB:AL111168; NID:G6968128; PIDN:CAB73196.1; PID:G6968376;  
 A:Experimental source: serotype O2, strain NCTC 11168  
 C:Genetics:

Gene: Cj0939c

Query Match 95.0%; Score 19; DB 2; Length 126;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|  
|  
|  
|  
b 83 VAEF 86

RESULT 39

70799

hypothetical protein RV3747 - Mycobacterium tuberculosis (strain H37RV)

Species: Mycobacterium tuberculosis  
Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
Accession: D70799  
Cole, S.T.; Broch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
Authors: Squares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
Reference number: A70500; MUID:98295987; PMID:9634230  
Accession: D70799  
Status: preliminary; nucleic acid sequence not shown; translation not shown  
Molecule type: DNA  
Residues: 1-127 <COL>  
Cross-references: GB:AL022121; GB:AL123456; NID:93261559; PIDN:CA118069.1; PID:9296017  
Experimental source: strain H37RV  
Genetics:  
Gene: RV3747

Query Match

Best Local Similarity 95.0%; Score 19; DB 2; Length 127;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|  
|  
|  
|  
b 80 VAEF 83

RESULT 40

AB1732

hypothetical protein lin2399 [imported] - Listeria innocua (strain Clip11262)

Species: Listeria innocua  
Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
Accession: AB1732  
Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma, K.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, W.; Title: Comparative genomics of Listeria species.  
Reference number: AB1077; MUID:21537279; PMID:11679669  
Accession: AB1732  
Status: preliminary  
Molecule type: DNA  
Residues: 1-127 <GLA>  
Cross-references: GB:AL592022; PIDN:CAC97626.1; PID:gi61414922; GSPDB:GN00178  
Experimental source: strain Clip11262  
Genetics:  
Gene: lin2399

Query Match

Best Local Similarity 95.0%; Score 19; DB 2; Length 127;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|  
|  
|  
|  
b 71 VAEF 74

RESULT 41

A54797

ileal lipid-binding protein - mouse  
Species: Mus musculus (house mouse)

Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 05-Nov-1999

Accession: A54797

Crossman, M.W.; Haft, S.M.; Gordon, J.I.

J. Cell Biol. 126, 1547-1564, 1994

Title: The mouse ileal lipid-binding protein gene: a model for studying axial pattern

Reference number: A54797; MUID:94375529; PMID:8089185

Accession: A54797

Status: preliminary

Molecule type: mRNA

Residues: 1-128 <CRO>

Cross-references: GB:U00938; NID:g507143; PIDN:AAC27352.1; PID:g507144

Genetics:

Introns: 23/1; 81/3; 111/3

Superfamily: myelin P2 protein

Query Match

Best Local Similarity 95.0%; Score 19; DB 2; Length 128;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

|  
|  
|  
|

Db 92 VAEF 95

RESULT 42

T02012

probable cathepsin B-like cysteine proteinase (EC 3.4.22.-) T15B16.17b - Arabidopsis thaliana

Species: Arabidopsis thaliana (mouse-ear cress)

Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 22-Jun-1999

Accession: T02012

Stonking, T.; Smith, R.

submitted to the EMBL Data Library, November 1998

Description: The sequence of A. thaliana T15B16.

Reference number: Z14488

Accession: T02012

Status: translated from GB/EMBL/DBJ

Molecule type: DNA

Residues: 1-129 <STO>

Cross-references: EMBL:AF104919; NID:g3859590; PIDN:AAC72873.1; PID:g3859607

Experimental source: cultivar Columbia

Genetics:

Map position: 4

Introns: 26/3; 44/3; 71/3

Note: T15B16.17b

Superfamily: papain

Keywords: cysteine proteinase; hydrolase

Query Match

Best Local Similarity 95.0%; Score 19; DB 2; Length 129;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

|  
|  
|  
|

Db 72 VAEF 75

RESULT 43

D72348

conserved hypothetical protein - Thermotoga maritima (strain MSB8)

Species: Thermotoga maritima

Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

Accession: D72348

Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey

Garratt, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999

Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

Reference number: A72200; MUID:99287316; PMID:10360571

Accession: D72348  
 Status: preliminary  
 Molecule type: DNA  
 Residues: 1-130 <ARN>  
 Cross-references: GB:AF001739; GB:AE000512; NID:G4981176; PID:RAD35739.1; PID:G498117  
 Experimental source: strain MSB8  
 Gene: TM0655  
 Superfamily: conserved hypothetical protein MW0315

Query Match 95.0%; Score 19; DB 2; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
 ||||  
 9 VAEF 12

RESULT 44  
 75208  
 hypothetical protein PAB2224 - Pyrococcus abyssi (strain Orsay)  
 Species: Pyrococcus abyssi  
 Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 06-Jan-2003  
 Accession: G75208  
 anonymous, Genoscope  
 Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome str  
 Reference number: A75001  
 Accession: G75208  
 Status: preliminary  
 Molecule type: DNA  
 Residues: 1-131 <NAW>  
 Cross-references: GB:AJ248283; GB:AL096836; NID:G5457433; PID: CAB49118.1; PID:G545762  
 Experimental source: strain Orsay  
 Geneticks:  
 Gene: PAB2224  
 Superfamily: uncharacterized conserved protein

Query Match 95.0%; Score 19; DB 2; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
 ||||  
 4 VAEF 7

RESULT 45  
 22270  
 hypothetical 15.2K protein - Anabaena sp.  
 Species: Anabaena sp.  
 Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 13-Sep-1998  
 Accession: JQ2270  
 Larimer, F.W.; Soper, T.S.  
 Title: Overproduction of Anabaena 7120 ribulose-bisphosphate carboxylase/oxygenase in  
 Reference number: JQ2270; MUID:93231541; PMID:8472962  
 Accession: JQ2270  
 Molecule type: DNA  
 Residues: 1-132 <LAR>  
 Cross-references: DDBJ:J01540  
 Experimental source: strain 7120  
 Geneticks:  
 Gene: rbcX  
 Superfamily: rbcX protein

Query Match 95.0%; Score 19; DB 2; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
 ||||

Db 81 VAEF 84

## RESULT 46

AG1996  
 hypothetical protein alr1525 [imported] - Nostoc sp. (strain PCC 7120)  
 Species: Nostoc sp. strain PCC 7120  
 Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 Accession: AG1996  
 R.Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anat  
 Reference number: AB1807; MUID:21595285; PMID:11759840  
 Accession: AG1996  
 Status: preliminary  
 Molecule type: DNA  
 Residues: 1-132 <KUR>  
 Cross-references: GB:BA000019; PID: BAB77891.1; PID:GL7135345; GSPDB:GN00179  
 Experimental source: strain PCC 7120  
 Geneticks:  
 Gene: alr1525  
 Superfamily: rbcX protein

Query Match 95.0%; Score 19; DB 2; Length 132;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

||||

Db 81 VAEF 84

## RESULT 47

AH0137  
 conserved hypothetical protein YP01120 [imported] - Yersinia pestis (strain CO92)  
 Species: Yersinia pestis  
 Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001  
 Accession: AH0137  
 R.Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; I  
 ll, M.; Rutherford, K.; Simmons, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, I  
 Nature 413, 523-527, 2001  
 Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 Reference number: AB0001; MUID:21470413; PMID:11586360  
 Accession: AH0137  
 Status: preliminary  
 Molecule type: DNA  
 Residues: 1-133 <KUR>  
 Cross-references: GB:AL590842; PID: CAC89963.1; PID:GL5979187; GSPDB:GN00175  
 Geneticks:  
 Gene: YP01120  
 Superfamily: 15.5K protein (tolAB operon 5' region)

Query Match 95.0%; Score 19; DB 2; Length 133;

Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

||||

Db 129 VAEF 132

## RESULT 48

H75623  
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
 Species: Deinococcus radiodurans  
 Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 Accession: H75623  
 R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; I  
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

```

Science 286, 1571-1577, 1999
;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
;Reference number: A75250; MUID:2003696; PMID:10567266
;Accession: H75623
;Status: preliminary
;Molecule type: DNA
;Residues: 1-133 <WHI>
;Cross-references: GB:AE001826; NID:g6460827; PIDN:AAF12590.1; PID:g6460886; TIGR:DRB00
;Experimental source: strain R1
;Genetics:
;Gene: DRB0052
;Map position: megaplasmid
;Genome: plasmid
;Note: plasmid MP1

Query Match 95.0%; Score 19; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
Db 18 VAEF 21

RESULT 49
riboflavin synthase beta subunit [imported] - Halobacterium sp. NRC-1
;Species: Halobacterium sp. NRC-1
;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
;Accession: A84221
;Authors: W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon
Jung, K.H.; Alam, M.; Freitas, T.
;Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
;Title: Genome sequence of Halobacterium species NRC-1.
;Reference number: A84160; MUID:20504483; PMID:11016950
;Accession: A84221
;Status: preliminary
;Molecule type: DNA
;Residues: 1-133 <STO>
;Cross-references: GB:AE004437; NID:g10580221; PIDN:AA819133.1; GSPDB:GN00138
;Genetics:
;Gene: ribE

Query Match 95.0%; Score 19; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
Db 8 VAEF 11

RESULT 50
ybcG protein - Escherichia coli (strain K-12)
;Species: Escherichia coli
;Date: 15-Oct-1982 #sequence_revision 30-Jun-1988 #text_change 01-Mar-2002
;Accession: A25980; G64809
;Sun, T.P.; Webster, R.E.
;Bacteriol. 169, 2667-2674, 1987
;Title: Nucleotide sequence of a gene cluster involved in entry of E. coli and single
;Reference number: A91835; MUID:87222192; PMID:3294803
;Accession: A25980
;Molecule type: DNA
;Residues: 1-134 <SUN>
;Cross-references: GB:M16489; NID:g148021; PIDN:AA83918.1; PID:g1128977
;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
;A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
;Title: The complete genome sequence of Escherichia coli K-12.
;Reference number: A64720; MUID:97426617; PMID:9278503

```

```

;Accession: G64809
;Status: nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-134 <BLAT>
;Cross-references: GB:AE000177; GB:U00096; NID:g1786955; PIDN:AA873830.1; PID:g1786957;
;Experimental source: strain K-12, substrain MGI655
;Comment: This is one of the proteins, encoded by the fli-tolAB gene cluster, that is i
;Genetics:
;Gene: ybcG
;Map position: 17 min
;Start codon: GTG
;Superfamily: 15.5K protein (tolAB operon 5' region)

```

```

Query Match 95.0%; Score 19; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
Db 129 VAEF 132

```

## RESULT 51

```

hypothetical protein ECs0771 [imported] - Escherichia coli (strain O157:H7, substrain RI
C90725
;Species: Escherichia coli
;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
;Accession: C90725
;Authors: T.; Makino, K.; Onishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
;Hayaishi, T.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
;Sawada, N.;
;DNA Res. 8, 11-24, 2001
;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
;Reference number: A99629; MUID:21156231; PMID:11258796
;Accession: C90725
;Status: preliminary
;Molecule type: DNA
;Residues: 1-134 <HAY>
;Cross-references: GB:BA000007; PIDN:BA834194.1; PID:g13360230; GSPDB:GN00154
;Experimental source: strain O157:H7, substrain RIMD 0509952
;Genetics:
;Gene: ECs0771
;Superfamily: 15.5K protein (tolAB operon 5' region)

```

```

Query Match 95.0%; Score 19; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
Db 129 VAEF 132

```

## RESULT 52

```

hypothetical protein ybcG [imported] - Escherichia coli (strain O157:H7, substrain EDL93
D85576
;Species: Escherichia coli
;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
;Accession: D85576
;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
;Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dinalanta, E.; Potamouisis, K.; Apodaca,
;Nature 409, 529-533, 2001
;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
;Reference number: A85480; MUID:21074935; PMID:11206551
;Accession: D85576
;Status: preliminary
;Molecule type: DNA
;Residues: 1-134 <STO>
;Cross-references: GB:AE005174; NID:g12513669; PIDN:AA855072.1; GSPDB:GN00145; UWGP:Z09
;Experimental source: strain O157:H7, substrain EDL933
;Genetics:
;Gene: ybcG
;Superfamily: 15.5K protein (tolAB operon 5' region)

```



Query Match 95.0%; Score 19; DB 2; Length 134;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
129 VAEF 132

SULT 53  
0592  
nerved hypothetical protein STY0790 [imported] - Salmonella enterica subsp. enterica  
Species: Salmonella enterica subsp. enterica serovar Typhi  
Note: this species has also been called Salmonella typhi  
Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
Accession: AD0592  
Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
S.; T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
ture 413, 848-852, 2001.  
Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
Reference number: AB0502; MUID:21534947; PMID:11677608  
Accession: AD0592  
Status: preliminary  
Molecule type: DNA  
Residues: 1-134 <PAR>  
Cross-references: GB:AL513382; PIDN:CAD05206.1; PID:gl6501976; GSPDB:GN00176  
Genetics:  
Gene: STY0790  
Superfamily: 15.5K protein (tolAB operon 5' region)

Query Match 95.0%; Score 19; DB 2; Length 134;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
129 VAEF 132

SULT 54  
6416  
pothetical protein - Synechocystis sp. (strain PCC 6803)  
Species: Synechocystis sp.  
Variety: PCC 6803  
Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
Accession: S76416  
Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
IA Res. 3, 109-136, 1996  
Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
Reference number: S74322; MUID:97061201; PMID:8905231  
Accession: S76416  
Status: preliminary  
Molecule type: DNA  
Residues: 1-136 <KAN>  
Cross-references: EMBL:D90915; GB:AB001339; NID:gl653604; PIDN:BA118545.1; PID:gl65363  
Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 95.0%; Score 19; DB 2; Length 136;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
88 VAEF 91

SULT 55  
11355  
pothetical protein Cj0829c [imported] - Campylobacter jejuni (strain NCTC 11168)

C:Species: Campylobacter jejuni  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
C:Accession: E81355  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrall  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: AB1250; MUID:20150912; PMID:10688204  
A:Accession: E81355  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-137 <PAR>  
A:Cross-references: GB:AL1139076; GB:AL111168; NID:G6968128; PIDN:CAB73094.1; PID:G6968271  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj0829c  
C:Superfamily: hypothetical protein ynfet

Query Match 95.0%; Score 19; DB 2; Length 137;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QV 2 VAEF 5  
Db 33 VAEF 36

RESULT 56  
F82106  
flagellar basal-body rod protein FlgC VC2199 [imported] - Vibrio cholerae (strain N16961  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: F82106  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
Chadson, B.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: F82106  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-138 <HEI>  
A:Cross-references: GB:AE004292; GB:AE003852; NID:G9656760; PIDN:AAF95344.1; GSPDB:GN001;  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC2199  
A:Map position: 1

Query Match 95.0%; Score 19; DB 2; Length 138;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QV 2 VAEF 5  
Db 79 VAEF 82

RESULT 57  
A69052  
riboflavin synthase beta subunit - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-May-2000  
C:Accession: A69052  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; E  
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;  
K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MUID:98037514; PMID:9371463  
A:Accession: A69052  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

```

;Residues: 1-139 <MTH>
;Cross-references: GB:AE000092; GB:AE000666; NID:g2622500; PIDN:AAB85867.1; PID:g2622500
;Experimental source: strain Delta H
;Genetics:
;Gene: MTH1390
;Superfamily: riboflavin synthase beta chain

Query Match      95.0%; Score 19; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

yY 2 VAEF 5
    ||||
bB 10 VAEF 13

RESULT 58
;Species: Bacillus halodurans
;Title: Hypothetical protein BH2601 [imported] - Bacillus halodurans (strain C-125)
;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
;Accession: AB3975
;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
;Nucleic Acids Res. 28, 4317-4331, 2000
;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
;Reference number: A83650; MUID:20512582; PMID:11058132
;Accession: AB3975
;Status: Preliminary
;Molecule type: DNA
;Residues: 1-139 <STO>
;Cross-references: GB:AF001516; GB:BA000004; NID:g10175192; PIDN:BA806320.1; GSPDB:GN00
;Experimental source: strain C-125
;Genetics:
;Gene: BH2601
;Superfamily: Bacillus subtilis hypothetical protein ylbD

Query Match      95.0%; Score 19; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

yY 2 VAEF 5
    ||||
bB 11 VAEF 14

RESULT 59
;Species: Archaeoglobus fulgidus
;Title: Hypothetical protein AF0586 - Archaeoglobus fulgidus
;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Nov-1999
;Accession: B69323
;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
;Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
;Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
;ature 390, 364-370, 1997
;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
;Smith, H.O.; Woese, C.R.; Venter, J.C.
;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
;Reference number: A69250; MUID:98049343; PMID:9389475
;Accession: B69323
;Status: preliminary; nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-140 <KLE>
;Cross-references: GB:AE001064; GB:AE000782; NID:g2689387; PIDN:AAB90658.1; PID:g265004
;Superfamily: Archaeoglobus fulgidus hypothetical protein AF0586

Query Match      95.0%; Score 19; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

yY 2 VAEF 5
    ||||
bB 65 VAEF 68

```

## RESULT 60

JQ2279

low-temperature regulated protein BN115 - rape

C;Species: Brassica napus (rape)

C;Date: 30-Sep-1993 #sequence\_revision 20-Aug-1994 #text\_change 28-May-1999

C;Accession: JQ2279

R;Weretilnyk, E.; Orr, W.; White, T.C.; Iu, B.; Singh, J.

Plant Physiol. 101, 171-177, 1993

A;Title: Characterization of three related low-temperature-regulated cDNAs from winter B

A;Reference number: JQ2279; MUID:94105287; PMID:7904076

A;Accession: JQ2279

A;Molecule type: mRNA

A;Residues: 1-142 &lt;WER&gt;

A;Cross-references: GB:S68726; NID:g544693; PIDN:AAB29482.1; PID:g544694

C;Comment: This protein is a low-temperature-related protein.

Query Match

Best Local Similarity 95.0%; Score 19; DB 2; Length 142;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5

Db 125 VAEF 128

## RESULT 61

JQ2280

low-temperature regulated protein BN19 - rape

C;Species: Brassica napus (rape)

C;Date: 30-Sep-1993 #sequence\_revision 20-Aug-1994 #text\_change 28-May-1999

C;Accession: JQ2280

R;Weretilnyk, E.; Orr, W.; White, T.C.; Iu, B.; Singh, J.

Plant Physiol. 101, 171-177, 1993

A;Title: Characterization of three related low-temperature-regulated cDNAs from winter B;

A;Reference number: JQ2279; MUID:94105287; PMID:7904076

A;Accession: JQ2280

A;Molecule type: mRNA

A;Residues: 1-142 &lt;WER&gt;

A;Cross-references: GB:S68879; NID:g544695; PIDN:AAB29483.1; PID:g544696

C;Comment: This protein is a low-temperature-related protein.

Query Match

Best Local Similarity 95.0%; Score 19; DB 2; Length 142;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5

Db 125 VAEF 128

## RESULT 62

H69515

riboflavin synthase, subunit beta (ribs) homolog - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 21-Jul-2000

C;Accession: H69515

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,

Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.;

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: H69515

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-143 &lt;KLE&gt;

A;Cross-references: GB:AE000957; GB:A3000782; NID:g2689280; PIDN:AAB89124.1; PID:g264840(

C;Superfamily: riboflavin synthase beta chain

```

Query Match          95.0%; Score 19; DB 2; Length 143;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

f      2 VAEF 5
      ||||
      10 VAEF 13

RESULT 63
36978
;Title: Streptomyces coelicolor (fragment)
;Species: Streptomyces coelicolor
;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
;Accession: T36978
;Author: K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
;Submitted to the EMBL Data Library, August 1999
;Reference number: Z21618
;Accession: T36978
;Status: preliminary; translated from GB/EMBL/DBJ
;Molecule type: DNA
;Residues: 1-143 <OLI>
;Cross-references: EMBL:AL109949; PIDN:CBAS2892.1; GSPDB:GN00070; SCOEDB:SCJ11.07c
;Experimental source: strain A3(2)
;Comment: A complete transposase sequence can be assembled from three adjacent orfs (PI
;Genetics:
;Gene: SCOEDB:SCJ11.07c

Query Match          95.0%; Score 19; DB 2; Length 143;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y      2 VAEF 5
      ||||
      113 VAEF 116

RESULT 64
41715
;Title: ribosomal protein L13 [similarity] - Haloarcula marismortui
;Species: Haloarcula marismortui
;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 14-Apr-2003
;Accession: B41715
;Author: Kroemer, W.J.; Arndt, E.
;Biochem. J. 266, 24573-24579, 1991
;Title: Halobacterial S9 operon. Three ribosomal protein genes are cotranscribed with g
;Genetics:
;Reference number: A41715; MUID:92105119; PMID:1840597
;Accession: B41715
;Status: preliminary
;Molecule type: DNA
;Residues: 1-145 <KRO>
;Cross-references: GB:M76567; NID:g148775; PIDN:AAA73097.1; PID:g148777
;Superfamily: ribosomal protein L13

Query Match          95.0%; Score 19; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y      2 VAEF 5
      ||||
      3 VAEF 6

RESULT 65
77043
;Title: hypothetical protein sll0678 - Synechocystis sp. (strain PCC 6803)
;Species: Synechocystis sp.
;Variety: PCC 6803
;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
;Accession: S77043
;Author: Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
;K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S77043
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-145 <KAN>
A:Cross-references: EMBL:D64005; GB:AB001339; NID:g1001779; PIDN:BAAL0735.1; PID:g100658;
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: Synechocystis hypothetical protein sll0678

Query Match          95.0%; Score 19; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY     2 VAEF 5
      ||||
      35 VAEF 38

Db

RESULT 66
F89837
;Title: conserved hypothetical protein SA0624 [imported] - Staphylococcus aureus (strain N315)
;Species: Staphylococcus aureus
;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
;Accession: F89837
;Author: R.; Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; I
;Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
;Cancr 357, 1285-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: X89758; MUID:21311952; PMID:11418146
A:Accession: F89837
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-146 <KUR>
A:Cross-references: GB:BA000018; PID:g13700560; PIDN:BAB41857.1; GSPDB:GN00149
A:Experimental source: strain N315
A:Genetics:
A:Gene: SA0624

Query Match          95.0%; Score 19; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY     2 VAEF 5
      ||||
      102 VAEF 105

Db

RESULT 67
C84294
;Title: hypothetical protein Vng1398c [imported] - Halobacterium sp. NRC-1
;Species: Halobacterium sp. NRC-1
;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
;Accession: C84294
;Author: R.; Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
;Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.
;Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: C84294
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-147 <STO>
A:Cross-references: GB:AE004437; NID:g10580903; PIDN:AAG19719.1; GSPDB:GN00138
A:Genetics:
A:Gene: VNG1398C
C:Superfamily: hypothetical protein yjcf

```

```

Query Match          95.0%; Score 19; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 2 VAEF 5
DB 117 VAEF 120

RESULT 68
T34803
Hypothetical protein SC2E1.35 - Streptomyces coelicolor
Species: Streptomyces coelicolor
Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 04-Mar-2000
Accession: T34803
Murphy, L.; Harris, D.; Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1998
Reference number: Z21557
Accession: T34803
Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: DNA
Residues: 1-147 <MUR>
Cross-references: EMBL:AL023797; PIDN:CAA19410.1; GSPDB:GN00070; SCOEDB:SC2E1.35
Experimental source: strain A3(2)
Genetics:
Superfamily: Streptomyces coelicolor hypothetical protein SC2E1.35

Query Match          95.0%; Score 19; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 2 VAEF 5
DB 125 VAEF 128

RESULT 69
B1788
Conserved hypothetical protein NMA2160 [imported] - Neisseria meningitidis (strain Z2491)
Species: Neisseria meningitidis
Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
Accession: B1788
Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
Hollroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
Reference number: B1775; MUID:20222556; PMID:10761919
Accession: B1788
Status: preliminary
Molecule type: DNA
Residues: 1-148 <PAR>
Cross-references: GB:AL157959; NID:G7380672; PIDN:CAB85372.1; PID:G738078
Experimental source: serogroup A, strain Z2491
Genetics:
Gene: NMA2160
Superfamily: hypothetical protein YTWI

Query Match          95.0%; Score 19; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 2 VAEF 5
DB 77 VAEF 80

RESULT 70
S10655
Hypothetical protein X - Pyrococcus woesei (fragment)
Species: Pyrococcus woesei
Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 18-Jun-1993
Accession: S10655

```

---

```

R;Zwickl, P.; Fabry, S.; Bogedain, C.; Haas, A.; Hensel, R.
J. Bacteriol. 172, 4329-4338, 1990
Title: Glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic archaeobacte
scherichia coli.
Reference number: S10650; MUID:90330536; PMID:2165475
Accession: S10655
Molecule type: DNA
Residues: 1-148 <ZWI>
Note: the authors translated the codon GGT for residue 54 as Glu

Query Match          95.0%; Score 19; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 2 VAEF 5
DB 115 VAEF 118

RESULT 71
AC1137
ribose 5-phosphate isomerase homolog lmo0498 [imported] - Listeria monocytogenes (strain
Species: Listeria monocytogenes
Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
Accession: AC1137
Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
Authors: Kreft, J.; Kuhn, M.; Kunst, P.; Kurapkak, G.; Madueno, E.; Maitournam, A.; Ma
ck, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boiland, J.A.; Voss, H.; Wehlend,
Title: Comparative genomics of Listeria species.
Reference number: AB1077; MUID:21537279; PMID:11679669
Accession: AC1137
Status: preliminary
Molecule type: DNA
Residues: 1-149 <GLA>
Cross-references: GB:NC_003210; PIDN:CAC98577.1; PID:G16409874; GSPDB:GN00177
Experimental source: strain EGD-e
Genetics:
Gene: lmo0498
Superfamily: galactoside O-acetyltransferase

Query Match          95.0%; Score 19; DB 2; Length 149;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 2 VAEF 5
DB 124 VAEF 127

RESULT 72
AB2857
pseudouridine [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
Species: Agrobacterium tumefaciens
Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
Accession: AB2857
Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCell
Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
scer, E.W.
Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
Reference number: AB2577; MUID:21608550; PMID:11743193
Accession: AB2857
Status: preliminary
Molecule type: DNA
Residues: 1-150 <KUR>
Cross-references: GB:AB008688; PIDN:AAL43272.1; PID:G17740759; GSPDB:GN00186
Experimental source: strain C58 (Dupont)
Genetics:

```

;Gene: Atu2283  
;Map position: circular chromosome  
;Superfamily: plastocyanin

Query Match 95.0%; Score 19; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
|||  
b 78 VAEF 81

# RESULT 73

BI089  
rotein gp35 from Bacteriophage A118 homolog lmo0113 [imported] - *Listeria monocytogenes*  
;Species: *Listeria monocytogenes*  
;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
;Accession: AB1089  
;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
; Jones, L.M.; Karst, U.  
; Science 294, 849-852, 2001  
; Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
; C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehländ,  
; Title: Comparative Genomics of *Listeria* species.  
; Reference number: AB1077; MUID:21537279; PMID:11679669  
; Accession: AB1089  
; Status: preliminary  
; Molecule type: DNA  
; Residues: 1-150 <GLA>  
; Cross-references: GB:NC\_003210; PID:NCAC98328.1; PID:gl6409472; GSPDB:GN00177  
; Experimental source: strain EGD-e  
; Genetics:  
; Gene: lmo0113

Query Match 95.0%; Score 19; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
|||  
b 105 VAEF 108

# RESULT 74

II452  
rotein gp35 from Bacteriophage A118 homolog lin0160 [imported] - *Listeria innocua* (stra  
;Species: *Listeria innocua*  
;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
;Accession: AI1452  
;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
; Jones, L.M.; Karst, U.  
; Science 294, 849-852, 2001  
; Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
; C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehländ,  
; Title: Comparative Genomics of *Listeria* species.  
; Reference number: AB1077; MUID:21537279; PMID:11679669  
; Accession: AI1452  
; Status: preliminary  
; Molecule type: DNA  
; Residues: 1-150 <GLA>  
; Cross-references: GB:AL592022; PID:NCAC95393.1; PID:gl6412579; GSPDB:GN00178  
; Experimental source: strain Clp11262  
; Genetics:  
; Gene: lin0160

Query Match 95.0%; Score 19; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5

Db 105 VAEF 108  
|||

# RESULT 75

H87263  
hypothetical protein CC0121 [imported] - *Caulobacter crescentus*  
;Species: *Caulobacter crescentus*  
;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
;Accession: H87263  
;Rieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.;  
; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
; J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
; Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
; A>Title: Complete Genome Sequence of *Caulobacter crescentus*.  
; A;Reference number: AB7249; MUID:21173698; PMID:11259647  
; A;Accession: H87263  
; A;Status: preliminary  
; A;Molecule type: DNA  
; A;Residues: 1-151 <STO>  
; A;Cross-references: GB:AE005673; NID:gl3421228; PID:AAK22108.1; GSPDB:GN00148  
; C;Genetics:  
; A;Gene: CC0121

Query Match 95.0%; Score 19; DB 2; Length 151;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
Db 73 VAEF 76

Search completed: May 24, 2004, 17:38:36  
Job time : 66.5 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model  
 run on: May 24, 2004, 17:33:28 ; Search time 8.57143 Seconds  
 (without alignments)  
 30.374 Million cell updates/sec

title: US-09-594-978A-1  
 effect score: 20  
 sequence: 1 XVAEF 5

coring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

searched: 141681 seqs, 52070155 residues

total number of hits satisfying chosen parameters: 141681

minimum DB seq length: 0  
 maximum DB seq length: 2000000000

post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 1000 summaries

database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	19	95.0	19	NS2_MYCTU	P81136 mycobacteri
2	19	95.0	19	SECE_TREPA	O83263 treponema p
3	19	95.0	59	YA90_ARCFU	O39175 archaeoglob
4	19	95.0	63	Y451_HAEIN	P43998 haemophilus
5	19	95.0	88	RS15_HAEIN	P44389 haemophilus
6	19	95.0	93	ALR2_METKA	O8twe6 methanopyru
7	19	95.0	108	YF24_METIA	O89919 methanococc
8	19	95.0	114	KEDA_ACTSL	P41249 actinomycet
9	19	95.0	115	INS_VERMO	Q9w7r2 verasper mo
10	19	95.0	116	INS_LOPPI	P01341 lophius pis
11	19	95.0	119	U279_DROME	Q9v8f3 drosophila
12	19	95.0	120	GTR2_PIG	O62786 sus scrofa
13	19	95.0	122	R18E_PYRAF	Q8zyq2 pyrobaculum
14	19	95.0	124	Y670_PASMU	Q8cm50 pasteurella
15	19	95.0	127	ILBP_MOUSE	P51162 mus musculu
16	19	95.0	130	SPEH_THEMA	Q9wzc3 thermotoga
17	19	95.0	131	Y194_PYRAB	Q9v280 pyrococcus
18	19	95.0	133	YBSC_HALN1	O8hrm5 halobacteri
19	19	95.0	134	YBSC_ECOLI	P08999 escherichia
20	19	95.0	139	R18E_WETH	O27443 methanobact
21	19	95.0	140	Y586_ARCFU	O39669 archaeoglob
22	19	95.0	143	R18B_ARCFU	O8152 archaeoglob
23	19	95.0	145	RL13_HALMA	P29198 haloarcula
24	19	95.0	148	CLM4_MOUSE	Q9jlm3 mus musculu
25	19	95.0	148	YORX_PYRWO	P20298 pyrococcus
26	19	95.0	150	R18B_PYRAF	O8zte3 pyrobaculum
27	19	95.0	152	SODC_PRIGL	P11418 prionace gl
28	19	95.0	154	R18B_SULSO	Q960B5 sulfolobus
29	19	95.0	155	PCP_FAEIN	P10325 haemophilus
30	19	95.0	157	RAP_TAROF	O49065 taraxacum o
31	19	95.0	157	R18E_SULTO	Q975m5 sulfolobus
32	19	95.0	163	SGS5_DROME	P07701 drosophila
33	19	95.0	166	UTR5_YEAST	P12630 saccharomyc

19	95.0	169	1	SSB_RHOBA	P59932 rhodopirell
19	95.0	172	1	BFL1_MOUSE	Q07440 mus musculu
19	95.0	175	1	BFL1_HUMAN	Q16548 homo sapien
19	95.0	176	1	FR12_RANCA	P07798 rana catesb
19	95.0	177	1	RL10_THETN	Q8r7u4 thermocanaer
19	95.0	178	1	DSBB_PASMU	Q516b3 pasteurella
19	95.0	178	1	ESMS_DROME	P13096 drosophila
19	95.0	178	1	YB18_ARCFU	O29147 archaeoglob
19	95.0	179	1	APT_HAEDU	Q7vkq4 haemophilus
19	95.0	183	1	RTTB_BOVIN	P18902 bos taurus
19	95.0	186	1	DNAI_WOLSP	P35907 wolbachia s
19	95.0	190	1	UCGI_PARDE	P05417 paracoccus
19	95.0	193	1	NTPA_METJA	Q57679 methanococc
19	95.0	194	1	HAM2_EACHD	Q9k8d9 bacillus na
19	95.0	194	1	TNR0_ECOLI	P05823 escherichia
19	95.0	196	1	PINQ_ECOLI	P77170 escherichia
19	95.0	196	1	PINR_ECOLI	P77574 escherichia
19	95.0	196	1	UHPA_ECOLI	P10940 escherichia
19	95.0	196	1	UHPA_SALTY	P27667 salmonella
19	95.0	198	1	AICD_MOUSE	Q3wve0 mus musculu
19	95.0	199	1	RHS5_SULSO	Q37777 sulfolobus
19	95.0	201	1	RETB_HORSE	Q28369 equus cabal
19	95.0	201	1	RETB_HUMAN	P02753 homo sapien
19	95.0	201	1	RETB_PIG	P27485 sus scrofa
19	95.0	201	1	RETB_RABIT	P06912 oryctolagus
19	95.0	204	1	YIGZ_ECOLI	P27862 escherichia
19	95.0	205	1	MD21_HUMAN	Q13257 homo sapien
19	95.0	205	1	MD21_MOUSE	Q21b5 mus musculu
19	95.0	208	1	YAC1_LEGPN	P37033 legionella
19	95.0	215	1	HNUO_CORDI	Q66120 cucumber mo
19	95.0	218	1	COAT_CMVIX	Q66120 cucumber mo
19	95.0	220	1	PTP1_NPVOP	O10274 orgyia pseu
19	95.0	223	1	ECR1_ARCFU	O29758 archaeoglob
19	95.0	223	1	RAN_TETPY	P41914 tetrahymena
19	95.0	223	1	VATD_TETH	O87880 thermus the
19	95.0	225	1	RAN_TETTH	P41915 tetrahymena
19	95.0	228	1	Y132_NPVOP	P24080 orgyia peeu
19	95.0	230	1	NADD_PSRPK	Q88d15 pseudomonas
19	95.0	230	1	RR3_FORPU	P51308 porphyra pu
19	95.0	230	1	UBIE_COREF	Q8f5b3 corynebacte
19	95.0	230	1	UBIE_CORGL	Q8nt39 corynebacte
19	95.0	234	1	YIGB_ECOLI	P24195 escherichia
19	95.0	236	1	PUR7_WOLSU	Q7m9x5 wolfinella s
19	95.0	236	1	YPS8_MYCTU	Q50740 mycobacteri
19	95.0	238	1	Y437_STAEP	O8ctk9 staphylococ
19	95.0	238	1	Y669_STAAM	O8nx12 staphylococ
19	95.0	238	1	YH89_STEMU	Q8ds14 streptococc
19	95.0	239	1	RLPA_NEIMA	Q9jnm7 neisseria m
19	95.0	239	1	RLPA_NEIMB	Q9k1a0 neisseria m
19	95.0	239	1	VG88_BPMLS	Q05305 mycobacteri
19	95.0	240	1	ATP7_ARATH	P58551 pasteurella
19	95.0	240	1	KDKA_PASPI	Q9ej12 arabidopsis
19	95.0	241	1	PCYA_PROMP	P58551 pasteurella
19	95.0	241	1	V755_AQURE	Q93tl5 prochloroco
19	95.0	242	1	COMB_SYNEL	O66957 aquifex aeo
19	95.0	244	1	ATPD_SORBI	Q87300 sorghum bic
19	95.0	247	1	CYSZ_VIBPA	Q87rj6 vibrio para
19	95.0	249	1	AQP_AEDAE	Q9nhw7 aedes aegyp
19	95.0	250	1	TIP1_TOBAC	P21653 nicotiana t
19	95.0	250	1	TIP2_TOBAC	P24422 nicotiana t
19	95.0	250	1	TIP3_TOBAC	P24422 nicotiana t
19	95.0	253	1	KDUD_ECOLI	P33560 attitrhinum
19	95.0	253	1	KDUD_ERWCH	Q05528 erwinia chr
19	95.0	255	1	PP1_BRANA	P23777 brassica na
19	95.0	255	1	TRMB_PASMU	Q9clc2 pasteurella
19	95.0	255	1	TRMB_HUMAN	P48556 homo sapien
19	95.0	257	1	PSD8_MOUSE	Q9cx56 mus musculu
19	95.0	257	1	PSD8_HUMAN	O28439 archaeoglob
19	95.0	258	1	NADC_ARCFU	O15400 homo sapien
19	95.0	261	1	STX7_HUMAN	O70439 mus musculu
19	95.0	261	1	STX7_MOUSE	O70257 rattus norv
19	95.0	266	1	TRPA_SYNEL	Q8dlng synechococc
19	95.0	267	1	APHC_MOUSE	Q9d099 mus musculu

107	19	95.0	267	1	DKGB_YERPE	Q8zh36 yersinia pe	180	19	95.0	325	1	UCP5_MOUSE	Q9z2b2 mus musculus
108	19	95.0	267	1	GNPI_CABEL	Q9xvj2 caenorhabdi	181	19	95.0	326	1	PE46_ARATH	O81772 arabidopsis
109	19	95.0	267	1	NW07_RAT	P50280 rattus norv	182	19	95.0	326	1	AIPL_CABEL	Q09444 caenorhabdi
110	19	95.0	269	1	AQ01_HUMAN	P29972 homo sapien	183	19	95.0	328	1	AIPL_RAT	Q911g9 rattus norv
111	19	95.0	269	1	AQ01_MOUSE	Q02013 mus musculu	184	19	95.0	328	1	Y778_AQUAE	O66974 aquifex aeo
112	19	95.0	269	1	RQ05_YEAST	P29975 rattus norv	185	19	95.0	329	1	YFV6_SCHPO	O927g8 schizosacch
113	19	95.0	269	1	RQ05_YEAST	P33334 saccharomyc	186	19	95.0	330	1	PRC2_FORGI	P33437 porphyromon
114	19	95.0	269	1	SRPB_MOUSE	P47758 mus musculu	187	19	95.0	330	1	Y840_METUA	Q58250 methanococc
115	19	95.0	269	1	YS13_STAAU	O05337 staphylococ	188	19	95.0	334	1	CCPA_BACSU	P25144 bacillus su
116	19	95.0	271	1	AQ01_BOVIN	P47865 bos taurus	189	19	95.0	334	1	PRC1_FORGI	P59916 porphyromon
117	19	95.0	271	1	SRPB_HUMAN	Q9v5m8 homo sapien	190	19	95.0	334	1	YJH3_RHIME	Q92v44 rhizobium m
118	19	95.0	272	1	AQ01_SHEEP	P56401 ovis aries	191	19	95.0	337	1	OPCM_CHICK	Q98892 gallus gall
119	19	95.0	272	1	YF63_STAMP	Q98rt3 staphylococ	192	19	95.0	338	1	O4XL_RHISN	P55704 rhizobium s
120	19	95.0	274	1	AMYC_THETU	P37729 thermansaer	193	19	95.0	339	1	HRDC_STROO	P18184 streptomyc
121	19	95.0	277	1	X087_THEMA	Q9cu22 thermotoga	194	19	95.0	339	1	IGB1_HUMAN	P78318 homo sapien
122	19	95.0	279	1	PHEA_LACLA	Q9ceuz lactococcus	195	19	95.0	340	1	HUR_STRAU	Q00923 streptococ
123	19	95.0	279	1	PHEA_LACLA	P43909 lactococcus	196	19	95.0	340	1	IGB1_MOUSE	Q61249 mus musculu
124	19	95.0	282	1	TRUEB_BORBU	O51743 borrelia bu	197	19	95.0	340	1	IGB1_RAT	O08836 rattus norv
125	19	95.0	284	1	Y3G4_SHEON	Q8aeaz2 shewanella	198	19	95.0	343	1	MO2M_ARATH	Q9m0m4 arabidopsis
126	19	95.0	284	1	YN13_MYCTU	P71899 mycobacteri	199	19	95.0	343	1	MO2M_ARATH	Q9f5k3 arabidopsis
127	19	95.0	285	1	YB46_HAEIN	P45071 haemophilus	200	19	95.0	343	1	XT30_ARATH	Q38908 arabidopsis
128	19	95.0	287	1	PDXV_ECOLI	P71150 escherichia	201	19	95.0	344	1	CD2_MOUSE	P08920 mus musculu
129	19	95.0	287	1	SYQ_DICDI	P14325 dictyostell	202	19	95.0	344	1	CD2_RAT	P08921 rattus norv
130	19	95.0	288	1	Y169_PASMU	Q9cp85 pasteurella	203	19	95.0	344	1	STSY_RAUSE	P15324 raunvolfia s
131	19	95.0	291	1	AMPM_ARCFU	O28438 archaeoglob	204	19	95.0	345	1	TRPD_SULSO	P50384 sulfolobus
132	19	95.0	292	1	YG29_METJA	Q59023 methanococ	205	19	95.0	347	1	FLIG_AGRTS	Q44458 agrobacteri
133	19	95.0	294	1	AMPM_METJA	Q58725 methanococ	206	19	95.0	352	1	CYCR_ROSDE	P26278 roseobacter
134	19	95.0	294	1	SSB1_YEAST	P10080 saccharomyc	207	19	95.0	352	1	STSY_CATRO	P18417 catharanthu
135	19	95.0	295	1	PRMA_VIECH	Q9kv64 vibrio chol	208	19	95.0	354	1	DCD1_HUMAN	P59894 homo sapien
136	19	95.0	295	1	RFBA_XANCP	P55256 xanthomonas	209	19	95.0	362	1	MURG_SHEON	Q8ck35 shewanella
137	19	95.0	295	1	YDH2_XANAU	P22644 xanthobacte	210	19	95.0	364	1	DDL_CHLTE	Q8kcr8 chlorobium
138	19	95.0	297	1	YD67_METJA	Q58762 methanococ	211	19	95.0	364	1	DHLE_BACLI	Q53560 bacillus li
139	19	95.0	299	1	T2R1_HUMAN	Q9nyw7 homo sapien	212	19	95.0	364	1	MRP_AQUAE	Q66946 aquifex aeo
140	19	95.0	300	1	NARE_CHICK	Q92080 gallus gall	213	19	95.0	365	1	MURG_SHEVI	Q9f1n0 shewanella
141	19	95.0	304	1	CYSD_XYLFA	Q9pd79 xyliella fas	214	19	95.0	366	1	CYCR_RHOGE	P51758 rhodocyclu
142	19	95.0	304	1	CYSD_XYLFA	Q87d98 xyliella fas	215	19	95.0	369	1	RPSD_STRPN	O03388 streptococ
143	19	95.0	305	1	E2BA_CABEL	P34604 caenorhabdi	216	19	95.0	378	1	KLAS_ECOLI	Q52328 escherichia
144	19	95.0	308	1	CPP3_ENTHI	Q06964 entamoeba h	217	19	95.0	379	1	TGT_VIECH	Q9kty9 vibrio chol
145	19	95.0	308	1	KSGA_BIFLO	Q8ge613 b dimethyla	218	19	95.0	380	1	METS_HELPY	Q92mw7 helicobacte
146	19	95.0	309	1	DNJL_MYCPN	Q50312 mycoplasma	219	19	95.0	380	1	METS_HELPY	P56069 helicobacte
147	19	95.0	309	1	NTCB_SVNY3	P74422 synecocyst	220	19	95.0	381	1	CD84_SCHPO	O09184 schizosacch
148	19	95.0	310	1	MIAA_CLOPE	Q84185 clostridium	221	19	95.0	383	1	CYCR_CHRVI	O82947 chromatium
149	19	95.0	310	1	PIR12_TOBAC	O04857 nicotiana t	222	19	95.0	384	1	AIPL_HUMAN	Q95mn3 pan paniscu
150	19	95.0	310	1	SIR5_HUMAN	Q9nx48 homo sapien	223	19	95.0	384	1	AIPL_FANPA	Q20701 caenorhabdi
151	19	95.0	310	1	SIR5_MOUSE	Q8k2c6 mus musculu	224	19	95.0	385	1	GBA5_CABEL	P43726 haemophilus
152	19	95.0	310	1	Y942_SULSO	Q97zh1 sulfolobus	225	19	95.0	385	1	PGK_HAEIN	P10312 bacterioph
153	19	95.0	311	1	DAPA_STRPN	Q97r25 streptococ	226	19	95.0	387	1	VPD_BPP2	P10312 bacterioph
154	19	95.0	311	1	DAPA_STRPN	Q8dp29 streptococ	227	19	95.0	388	1	METS_MYCLE	P46807 mycobacteri
155	19	95.0	311	1	MIAA_LACPL	Q88wp5 lactobacill	228	19	95.0	388	1	METS_MYCTU	O53427 mycobacteri
156	19	95.0	311	1	OSL1_HUMAN	Q8ng12 homo sapien	229	19	95.0	389	1	RFIM_CABEL	O44568 caenorhabdi
157	19	95.0	311	1	OSL2_HUMAN	Q8ng10 homo sapien	230	19	95.0	389	1	VPD_BF186	P21679 bacterioph
158	19	95.0	312	1	O6N1_HUMAN	Q8ng55 homo sapien	231	19	95.0	390	1	Y181_ARCFU	O30056 archaeoglob
159	19	95.0	312	1	PP12_ARATH	P48482 arabidopsis	232	19	95.0	392	1	TEEB_STYMY	P29548 stylonichia
160	19	95.0	312	1	PP15_ARATH	P48485 arabidopsis	233	19	95.0	392	1	YBVF_HAEIN	P44083 haemophilus
161	19	95.0	313	1	DRPE_CRAPL	P22242 cratersostig	234	19	95.0	394	1	DXR_SYNY3	Q55663 synecocyst
162	19	95.0	313	1	PEPM_STRHY	P29247 streptomyc	235	19	95.0	395	1	CGA2_CHICK	P43449 gallus gall
163	19	95.0	314	1	LUNE_CRIFA	Q2546 crithidia f	236	19	95.0	397	1	Y929_THEMA	P56727 thermotoga
164	19	95.0	314	1	PP1_FHAVU	P48490 phaseolus v	237	19	95.0	398	1	PNCB_BUCAP	Q8k9i6 buchnera ap
165	19	95.0	315	1	O3A2_HUMAN	P47893 homo sapien	238	19	95.0	399	1	PNCB_BUCAL	P57442 buchnera ap
166	19	95.0	315	1	O3A2_PANTR	Q9tu97 pan troglod	239	19	95.0	400	1	ER19_HUMAN	P53602 homo sapien
167	19	95.0	315	1	O3A3_HUMAN	P47888 homo sapien	240	19	95.0	401	1	ER19_MOUSE	Q99jfs mus musculu
168	19	95.0	315	1	O3A3_PANTR	Q9tua0 pan troglod	241	19	95.0	401	1	ER19_RAT	Q62967 rattus norv
169	19	95.0	316	1	APEI_MOUSE	P28352 mus musculu	242	19	95.0	402	1	C123_MYCTU	P77902 mycobacteri
170	19	95.0	316	1	APEI_RAT	P43138 rattus norv	243	19	95.0	402	1	C13B_XYLFA	Q9pgcs xyliella fas
171	19	95.0	316	1	PP1_BRAOL	P48487 brassica ol	244	19	95.0	402	1	C13B_XYLFT	O87ax5 xyliella fas
172	19	95.0	317	1	APEI_HUMAN	P27695 homo sapien	245	19	95.0	406	1	CGA2_BOVIN	P30274 bos taurus
173	19	95.0	317	1	OSTP_RAT	P08721 rattus norv	246	19	95.0	406	1	T230_HUMAN	P48775 homo sapien
174	19	95.0	317	1	PP11_TOBAC	O04856 nicotiana t	247	19	95.0	407	1	RPF2_HUMAN	O60858 homo sapien
175	19	95.0	318	1	PP11_ARATH	P30366 arabidopsis	248	19	95.0	407	1	Y4DM_RHISN	P55412 rhizobium s
176	19	95.0	322	1	PP16_ARATH	P48486 arabidopsis	249	19	95.0	409	1	Y4DM_RHISN	Q14186 homo sapien
177	19	95.0	324	1	ADD_RHIME	Q92r48 rhizobium m	250	19	95.0	410	1	TDPI_HUMAN	Q08639 mus musculu
178	19	95.0	324	1	PP18_ARATH	O82734 arabidopsis	251	19	95.0	411	1	PBPT_YERPE	Q8zfr0 yersinia pe
179	19	95.0	325	1	UCP5_HUMAN	Q95258 homo sapien	252	19	95.0	412	1	CCA_ECOLI	P06961 escherichia

253	19	95.0	415	1	CGA2_XENLA	P47827 xenopus lae	326	19	95.0	477	1	CAP2_HUMAN	P40123 homo sapien
254	19	95.0	415	1	HEM1_NEIGA	Q2hd36 neisseria g	327	19	95.0	477	1	CAP2_RAT	P52481 rattus norv
255	19	95.0	415	1	HEM1_NEIBA	Q2hd35 neisseria m	328	19	95.0	479	1	RFBM_SALTY	P26404 salmonella
256	19	95.0	415	1	HEM1_NEIMB	P56994 neisseria m	329	19	95.0	479	1	YSO2_CAEEL	Q09357 caenorhabdi
257	19	95.0	417	1	ASSY_OCEHL	Q8elt8 oceanobacil	330	19	95.0	483	1	NIFE_METMP	P71528 methanococc
258	19	95.0	417	1	HEM1_CHRVO	P56173 chromobacte	331	19	95.0	486	1	GLG2_RHIME	P58394 rhizobium m
259	19	95.0	418	1	CSD_STRCO	Q9xads streptomyc	332	19	95.0	488	1	ARL2_RHIME	Q92vms rhizobium m
260	19	95.0	419	1	MURD_CHLVC	Q82161 chlamydophi	333	19	95.0	490	1	ARLY_BIFLO	Q8G5f3 bifidobacte
261	19	95.0	420	1	CGA2_CHLVR	P15986 chlorohydra	334	19	95.0	491	1	AP2A_SHEEP	Q9n0n3 ovis aries
262	19	95.0	420	1	CGA2_MESAU	P37881 mesocricetu	335	19	95.0	491	1	G6PD_ERWCH	P37986 erwina chr
263	19	95.0	421	1	MUAI_STAMM	Q398f8 staphylococ	336	19	95.0	491	1	GCSB_COXBU	Q83b09 coxiella bu
264	19	95.0	421	1	MUAI_STAPB	Q8crn8 staphylococ	337	19	95.0	491	1	NU4M_METSE	Q84797 metridium s
265	19	95.0	422	1	CGA2_MOUSE	P51943 mus musculu	338	19	95.0	492	1	FTSA_HELPF	O25629 helicobacte
266	19	95.0	422	1	HEM1_PSEAE	P42848 thermotoga	339	19	95.0	493	1	FTSA_HELPF	Q9zkm3 helicobacte
267	19	95.0	423	1	CISY_ACIAN	P20902 acetobact	340	19	95.0	493	1	GLYM_CANAL	O13425 candida alb
268	19	95.0	423	1	GLYA_METH	O77433 methanobact	341	19	95.0	494	1	PRE_STRAG	P13925 streptococc
269	19	95.0	423	1	GRB4_CAEEL	Q09453 caenorhabdi	342	19	95.0	498	1	NU4C_PSINU	Q8whx8 psilotom nu
270	19	95.0	424	1	GLA2_BALSO	Q8xt41 ralstonia s	343	19	95.0	499	1	PTFA_HASIN	P44715 h pts syste
271	19	95.0	426	1	RHO_DEIRA	P52153 deinococcus	344	19	95.0	500	1	CATA_NEIGO	O59602 neisseria g
272	19	95.0	429	1	ENO_THEMA	P42848 thermotoga	345	19	95.0	500	1	DNAA_BIFLO	Q896K0 bifidobacte
273	19	95.0	430	1	SVH_CHLBE	Q8kft6 chlorobium	346	19	95.0	503	1	AMPA_RHILO	Q98481 rhizobium l
274	19	95.0	431	1	ENO_CLOTE	Q97152 clostridium	347	19	95.0	503	1	UBIB_NEIMA	Q9jvd5 neisseria m
275	19	95.0	431	1	HNDH_BORBU	O51628 borrelia bu	348	19	95.0	503	1	UBIB_NEIMB	Q9K0n0 neisseria m
276	19	95.0	432	1	CGA2_HUMAN	P20248 homo sapien	349	19	95.0	504	1	DCMC_ANSAN	P12617 anser anser
277	19	95.0	432	1	GLYA_BRAJA	P24060 bradyrhizob	350	19	95.0	504	1	YO93_CAEEL	P41843 caenorhabdi
278	19	95.0	433	1	C125_MYCTU	P71856 mycobacteri	351	19	95.0	506	1	PUR9_AQUAE	O67775 a bifunctio
279	19	95.0	433	1	ENGA_AQUAE	O67749 aquifex aeo	352	19	95.0	512	1	C981_SORBI	O48956 sorghum bic
280	19	95.0	434	1	GLYA_HYPEE	P34895 mychomicrob	353	19	95.0	513	1	NHAB_ECOLI	P27377 escherichia
281	19	95.0	434	1	GLYA_METEX	P50435 methylobact	354	19	95.0	514	1	YHK8_YEAST	P38776 saccharomyc
282	19	95.0	436	1	HNDH_ARCTU	O28538 archaeoglob	355	19	95.0	518	1	AMT6_BAGS7	P19571 bacillus sp
283	19	95.0	437	1	AP2A_HUMAN	P55549 homo sapien	356	19	95.0	520	1	BOHB_HEIMO	Q92958 helicobacill
284	19	95.0	437	1	AP2A_MOUSE	P34056 mus musculu	357	19	95.0	520	1	DHAC_PSRPU	P09788 pseudomonas
285	19	95.0	437	1	AP2A_RAT	P56197 rattus norv	358	19	95.0	521	1	SYE_LEPIN	O8eym3 leptospira
286	19	95.0	437	1	FOLC_HABIN	P43775 haemophilus	359	19	95.0	522	1	GR22_RAT	P12336 rattus norv
287	19	95.0	437	1	GNTT_ECOLI	P39835 escherichia	360	19	95.0	522	1	NHAB_HABIN	P44706 haemophilus
288	19	95.0	438	1	G3PB_TORAB	P09044 nicotiana t	361	19	95.0	523	1	GTR2_MOUSE	P14246 mus musculu
289	19	95.0	438	1	GLYA_BRUME	Q8ygg7 brucella me	362	19	95.0	524	1	IMDH_SCHPO	P11168 homo sapien
290	19	95.0	440	1	GLYA_BRUSU	Q95118 brucella su	363	19	95.0	524	1	IMDH_SCHPO	P11168 homo sapien
291	19	95.0	445	1	NHRI_CAEEL	Q18187 caenorhabdi	364	19	95.0	525	1	UBIB_RALSO	O14344 schizosacch
292	19	95.0	445	1	ZGS_XENLA	P18725 xenopus lae	365	19	95.0	530	1	RNP2_HUMAN	O8y275 ralstonia s
293	19	95.0	446	1	DTA2_RALSO	Q8xub6 ralstonia s	366	19	95.0	530	1	RNP2_MOUSE	O1498 homo sapien
294	19	95.0	446	1	SYE_AZOB	P45631 azospirillum	367	19	95.0	532	1	PHR_FOTTR	O8vhl1 mus musculu
295	19	95.0	447	1	G3PE_ARATH	P5857 arabidopsis	368	19	95.0	536	1	PUR9_RHIME	Q28811 potoruso tr
296	19	95.0	448	1	YX43_CAEEL	P34255 caenorhabdi	369	19	95.0	537	1	4CL_PINTA	Q92Kx6 r bifunctio
297	19	95.0	449	1	G64F_DROME	P32327 drosophila	370	19	95.0	538	1	PUR9_AGR15	P41636 pinus taeda
298	19	95.0	449	1	GABT_MYCTU	Q50632 mycobacteri	371	19	95.0	538	1	PUR9_BRUME	Q8ubm8 a bifunctio
299	19	95.0	450	1	AMYG_SCHPO	Q60087 schizosacch	372	19	95.0	539	1	Z173_HUMAN	O8y153 b bifunctio
300	19	95.0	450	1	SYN_ENTFA	Q831x4 enterococcu	373	19	95.0	540	1	VIBE_VIBCH	O12899 homo sapien
301	19	95.0	453	1	YKX8_YEAST	Q9er04 mus musculu	374	19	95.0	543	1	G6PI_CORGL	Q7yr34 pan troglod
302	19	95.0	455	1	TKS5_MOUSE	P57909 pasteurella	375	19	95.0	547	1	VIBE_VIBCH	Q07899 vibrio chol
303	19	95.0	457	1	ARLY_PASMU	Q9h343 homo sapien	376	19	95.0	548	1	TCPA_TETPY	O15891 tetrahymena
304	19	95.0	457	1	TKS5_HUMAN	Q9h343 homo sapien	377	19	95.0	548	1	G6PI_NEIMA	Q9jtw1 neisseria m
305	19	95.0	458	1	Z275_HUMAN	Q9ns44 homo sapien	378	19	95.0	548	1	G6PI_NEIMB	Q9jyx3 neisseria m
306	19	95.0	459	1	REPS_STRAM	P36891 streptomyc	379	19	95.0	548	1	PPAC_THEMEA	Q9wz56 thermotoga
307	19	95.0	460	1	FUMC_CHLVC	Q822d5 chlamydophi	380	19	95.0	549	1	G6PI_PASMU	Q9cn12 pasteurella
308	19	95.0	461	1	ARLY_RAT	P20673 rattus norv	381	19	95.0	549	1	RSI_HABIN	Q48082 haemophilus
309	19	95.0	461	1	FUMC_RICPR	Q9zcg4 rickettsia	382	19	95.0	549	1	TCPA_CAEEL	P41988 caenorhabdi
310	19	95.0	463	1	FUMC_RICCN	Q92gw0 rickettsia	383	19	95.0	550	1	G6PI_VIBCH	Q9kuy4 vibrio chol
311	19	95.0	464	1	ARLY_HUMAN	P04424 homo sapien	384	19	95.0	550	1	YBYC_CAEEL	Q21802 caenorhabdi
312	19	95.0	464	1	ARLY_MOUSE	Q1y140 mus musculu	385	19	95.0	553	1	HCP_VIBU	Q8d8v3 vibrio vuln
313	19	95.0	465	1	SELA_DESBA	P56372 desulfovibr	386	19	95.0	553	1	NIFA_RHOCA	P09434 rhodobacter
314	19	95.0	468	1	P2C3_YEAST	P34221 saccharomyc	387	19	95.0	553	1	SYR_STAPF	Q8ctn9 staphylococ
315	19	95.0	469	1	MM01_HORSE	Q9xsz5 equus cabal	388	19	95.0	555	1	MASY_PICAN	P21360 pichia angu
316	19	95.0	469	1	MM01_PIG	P21692 sus scrofa	389	19	95.0	574	1	TP6B_AERPE	Q9y64 aeropyrum p
317	19	95.0	471	1	MANC_SALMO	Q1410 salmonella	390	19	95.0	574	1	PT1_LACSK	O07126 lactobacill
318	19	95.0	473	1	ARLY_STRLC	P50988 streptomyc	391	19	95.0	575	1	MBHL_WOLSU	P31883 wolinnella s
319	19	95.0	473	1	RFBM_SALMU	Q00473 salmonella	392	19	95.0	575	1	MSPA_TREMA	Q9z413 treponema m
320	19	95.0	474	1	DCDA_MYCSM	Q9x5m1 mycobacteri	393	19	95.0	576	1	DD34_HUMAN	Q14147 homo sapien
321	19	95.0	475	1	ARLY_STRCO	Q911b1 streptomyc	394	19	95.0	576	1	SYP_HELPD	Q9zmi6 helicobacte
322	19	95.0	475	1	ECHE_RAT	Q50587 rattus norv	395	19	95.0	577	1	PT1_STRBO	Q9wxk3 streptococc
323	19	95.0	476	1	CAP2_MOUSE	Q50587 rattus norv	396	19	95.0	577	1	SYP_HELPF	P56124 helicobacte
324	19	95.0	476	1	GLGA_HABIN	P45179 haemophilus	397	19	95.0	578	1	V022_FOWPV	Q9j5h9 fowpox vir
325	19	95.0	476	1	NOEA_RHIME	Q28892 rhizobium m	398	19	95.0	579	1	FUMA_SALTY	P40720 salmonella



399	19	95.0	580	1	RS1_CHLPH	Q928m3 chlamydia p	472	19	95.0	649	1	INVA_PHAU	P29001 phaseolus a
400	19	95.0	587	1	CO8B_ONCMY	Q90x85 oncorhynchu	473	19	95.0	651	1	INVA_PHAU	Q24509 phaseolus v
401	19	95.0	588	1	RS1_XENLA	P13563 xenopus lae	474	19	95.0	652	1	C3BB_BACTU	Q06117 bacillus th
402	19	95.0	589	1	GBP1_MOUSE	Q01514 mus musculus	475	19	95.0	660	1	FDL_DROME	Q8w5f3 drosophila
403	19	95.0	589	1	GBP2_MOUSE	Q920e6 mus musculus	476	19	95.0	660	1	TP6B_METJA	Q58434 methanococc
404	19	95.0	589	1	GBP2_MOUSE	Q920e6 mus musculus	477	19	95.0	670	1	NICA_ARATH	Q8gum5 arabidopsis
405	19	95.0	591	1	GBP2_MOUSE	Q920e6 mus musculus	478	19	95.0	677	1	NCPR_MOUSE	Q37040 mus musculus
406	19	95.0	592	1	ALAM_YEAST	P32456 homo sapien	479	19	95.0	677	1	NCPR_MOUSE	P00388 rattus norv
407	19	95.0	592	1	ALAM_YEAST	P32456 homo sapien	480	19	95.0	680	1	DCP_SALTY	P27236 salmonella
408	19	95.0	602	1	PEX5_HUMAN	P50342 homo sapien	481	19	95.0	685	1	CSD_MYCPA	Q9x116 mycobacteri
409	19	95.0	604	1	TP6B_ARCFU	P29605 archaeoglob	482	19	95.0	690	1	TRF1_SALSA	P80426 salmo salar
410	19	95.0	604	1	DNAB_BACME	P05646 bacillus me	483	19	95.0	691	1	TRF2_SALSA	P80429 salmo salar
411	19	95.0	605	1	APM2_MOUSE	P38700 saccharomyc	484	19	95.0	692	1	HEXA_ANOGA	Q17020 anopheles g
412	19	95.0	607	1	BKL3_YEAST	Q8bhi4 mus musculus	485	19	95.0	693	1	RECG_PASMU	Q9cmb4 pasteurella
413	19	95.0	607	1	DNAB_FUSNM	Q8rh05 fusobacteri	486	19	95.0	695	1	TRPE_PABIT	P19134 cryptolagus
414	19	95.0	607	1	DNAB_LACLA	P42168 lactococcus	487	19	95.0	696	1	TRPE_PABIT	P19134 cryptolagus
415	19	95.0	607	1	DNAB_STRP3	Q8k624 streptococc	488	19	95.0	698	1	TRFE_HUMAN	P03571 sus scrofa
416	19	95.0	607	1	DNAB_STRPN	P95829 streptococc	489	19	95.0	704	1	ICA_FIG	P02787 homo sapien
417	19	95.0	607	1	DNAB_STRPN	P95831 streptococc	490	19	95.0	704	1	TRFE_HUMAN	Q29545 sus scrofa
418	19	95.0	607	1	DNAB_STRPN	P95831 streptococc	491	19	95.0	705	1	TRFE_HUMAN	Q29443 bos taurus
419	19	95.0	607	1	DNAB_STRPN	P95831 streptococc	492	19	95.0	705	1	TRFE_HUMAN	Q84474 bacteroides
420	19	95.0	609	1	BKL3_HUMAN	Q8ab22 homo sapien	493	19	95.0	706	1	TRFE_HUMAN	Q49836 lithospermu
421	19	95.0	610	1	DNAB_ENTFA	Q835r7 enterococc	494	19	95.0	706	1	TRFE_HUMAN	Q9x86 mus musculus
422	19	95.0	610	1	DNAB_BACSH	Q69268 bacillus sp	495	19	95.0	706	1	TRFE_HUMAN	Q85983 leptospira
423	19	95.0	611	1	DNAB_STRPN	Q99tr7 staphylococ	496	19	95.0	707	1	TRFE_HUMAN	P27425 equus cabal
424	19	95.0	611	1	DNAB_BACAA	P45554 staphylococ	497	19	95.0	707	1	TRFE_HUMAN	Q07266 rattus norv
425	19	95.0	611	1	DNAB_BACCR	Q811s2 bacillus an	498	19	95.0	708	1	TRFE_HUMAN	P14166 ipomoea bat
426	19	95.0	612	1	DNAB_LISIN	Q818e9 bacillus ce	499	19	95.0	708	1	TRFE_HUMAN	Q23865 daucus caro
427	19	95.0	612	1	DNAB_LISMO	Q92bn8 listeria in	500	19	95.0	708	1	TRFE_HUMAN	Q42858 ipomoea bat
428	19	95.0	612	1	ILVD_XANCP	Q955a4 listeria mo	501	19	95.0	710	1	TRFE_HUMAN	P03070 simian viru
429	19	95.0	612	1	ILVD_XANCP	Q8p9i0 xanthomonas	502	19	95.0	712	1	TRFE_HUMAN	Q49835 lithospermu
430	19	95.0	614	1	YC81_MYCTU	Q8p9j3 xanthomonas	503	19	95.0	712	1	TRFE_HUMAN	P35513 nicotiana t
431	19	95.0	614	1	DNAB_LACSN	Q11040 mycobacteri	504	19	95.0	714	1	TRFE_HUMAN	P45733 nicotiana t
432	19	95.0	621	1	SPKE_SYNY3	Q8km16 lactobacill	505	19	95.0	715	1	TRFE_HUMAN	Q44219 drosophila
433	19	95.0	621	1	TP6B_METFA	P73515 synchocyst	506	19	95.0	716	1	TRFE_HUMAN	P28793 p fatty oxi
434	19	95.0	622	1	TP6B_METFA	Q8tcf7 methanosarc	507	19	95.0	721	1	TRFE_HUMAN	Q8ypt0 anabaena sp
435	19	95.0	622	1	TP6B_METFA	Q8p9b8 methanosarc	508	19	95.0	722	1	TRFE_HUMAN	Q10767 mycobacteri
436	19	95.0	624	1	COOS_METJA	Q45743 bacillus th	509	19	95.0	725	1	TRFE_HUMAN	Q55690 synchocyst
437	19	95.0	624	1	P2L2_HUMAN	Q58138 methanococc	510	19	95.0	725	1	TRFE_HUMAN	P35510 arabidopsis
438	19	95.0	625	1	TP22_HAEN	Q9nm6 homo sapien	511	19	95.0	729	1	TRFE_HUMAN	P35510 arabidopsis
439	19	95.0	627	1	TP22_HAEN	P44971 haemophilus	512	19	95.0	759	1	TRFE_HUMAN	P23842 escherichia
440	19	95.0	627	1	CACP_CANTR	Q00614 candida tro	513	19	95.0	763	1	TRFE_HUMAN	Q8kd17 chlorobium
441	19	95.0	631	1	CACP_CANTR	P52826 columba liv	514	19	95.0	764	1	TRFE_HUMAN	P31075 wolinsella s
442	19	95.0	631	1	NTPL1_VACCC	Q57214 vaccinia vi	515	19	95.0	764	1	TRFE_HUMAN	Q42669 cucumis mel
443	19	95.0	631	1	NTPL1_VACCV	P20637 vaccinia vi	516	19	95.0	781	1	TRFE_HUMAN	Q29753 archaeoglob
444	19	95.0	631	1	NTPL1_VACCV	P05807 vaccinia vi	517	19	95.0	787	1	TRFE_HUMAN	Q29753 archaeoglob
445	19	95.0	632	1	NTPL1_MYXVL	P33066 variola vir	518	19	95.0	787	1	TRFE_HUMAN	Q84253 chlamydia t
446	19	95.0	632	1	NTPL1_SFVKA	Q9q814 myxoma viru	519	19	95.0	804	1	TRFE_HUMAN	Q10313 schizosacch
447	19	95.0	634	1	DNAB_HASIN	Q9g822 shope fibro	520	19	95.0	807	1	TRFE_HUMAN	Q9hr31 halobacteri
448	19	95.0	634	1	NTPL1_MCVI	P43736 haemophilus	521	19	95.0	807	1	TRFE_HUMAN	Q91w96 mus musculus
449	19	95.0	635	1	DNAB_VIRGH	Q98267 molluscum c	522	19	95.0	808	1	TRFE_HUMAN	Q9ujx5 homo sapien
450	19	95.0	635	1	DNAB_VIRGH	Q34241 vibrio chol	523	19	95.0	808	1	TRFE_HUMAN	Q9ujx5 homo sapien
451	19	95.0	636	1	DNAB_VIRGH	Q8y280 anabaena sp	524	19	95.0	808	1	TRFE_HUMAN	Q14255 schizosacch
452	19	95.0	636	1	DNAB_VIRGH	Q8d566 vibrio vuln	525	19	95.0	811	1	TRFE_HUMAN	Q56239 thermus the
453	19	95.0	636	1	MYB_MOUSE	Q7m885 vibrio vuln	526	19	95.0	821	1	TRFE_HUMAN	Q8dl37 synchococc
454	19	95.0	637	1	DNAB_BRUME	P06876 mus musculus	527	19	95.0	821	1	TRFE_HUMAN	P14346 herpesviru
455	19	95.0	637	1	DNAB_BRUOV	Q8ye76 bruceella me	528	19	95.0	842	1	TRFE_HUMAN	P43533 pichia past
456	19	95.0	637	1	DNAB_BRUSU	Q05981 bruceella ov	529	19	95.0	858	1	TRFE_HUMAN	P53532 corynebacte
457	19	95.0	637	1	DNAB_ECOLI	Q8fx22 bruceella gu	530	19	95.0	858	1	TRFE_HUMAN	P30594 rhizopus ol
458	19	95.0	637	1	DNAB_PSEAE	P04475 escherichia	531	19	95.0	861	1	TRFE_HUMAN	P34552 caenorhabdi
459	19	95.0	637	1	DNAB_SALTI	Q9hw43 pseudomonas	532	19	95.0	861	1	TRFE_HUMAN	Q51737 borrelia bu
460	19	95.0	637	1	DNAB_SALTI	Q8z9r1 salmonella	533	19	95.0	876	1	TRFE_HUMAN	Q19954 caenorhabdi
461	19	95.0	637	1	DNAB_SALTY	Q56073 salmonella	534	19	95.0	876	1	TRFE_HUMAN	Q9cdsl lactococcus
462	19	95.0	637	1	DNAB_VIRPA	Q87rx3 vibrio para	535	19	95.0	877	1	TRFE_HUMAN	Q32801 lactococcus
463	19	95.0	638	1	DNAB_VIRBP	Q91721 vibrio prot	536	19	95.0	877	1	TRFE_HUMAN	P59199 streptococ
464	19	95.0	638	1	DNAB_BUCBP	P59565 buchnera ap	537	19	95.0	877	1	TRFE_HUMAN	P28277 herpes simp
465	19	95.0	638	1	DNAB_RHILE	Q33228 rhizobium l	538	19	95.0	881	1	TRFE_HUMAN	Q8d556 synchococc
466	19	95.0	638	1	DNAB_SHIFL	Q83mns shigella fl	539	19	95.0	882	1	TRFE_HUMAN	Q9un79 homo sapien
467	19	95.0	640	1	DNAB_VIBHA	Q87384 vibrio harv	540	19	95.0	889	1	TRFE_HUMAN	Q9un79 homo sapien
468	19	95.0	640	1	MYB_BOVIN	P46200 bos taurus	541	19	95.0	894	1	TRFE_HUMAN	Q8r3s6 mus musculus
469	19	95.0	640	1	MYB_BOVIN	Q12422 homo sapien	542	19	95.0	895	1	TRFE_HUMAN	P28877 candida alb
470	19	95.0	642	1	INVA_VICFA	Q43857 vicia faba	543	19	95.0	899	1	TRFE_HUMAN	Q04500 saccharomyc
471	19	95.0	644	1	H570_BRUMA	P27541 bugia maia	544	19	95.0	916	1	TRFE_HUMAN	Q07421 ajellomyces
	19	95.0	644	1	LSU2_RHIFA	P17279 rhizomucor							
	19	95.0	649	1	DREB_HUMAN	Q16643 homo sapien							

545	19	95.0	916	1	SVI_STARP	O8cxs1	staphylococ	618	19	95.0	2236	1	POLG_HPAV4	P26581	hepatitis a
546	19	95.0	919	1	Y893_HUMAN	O94967	homo sapien	619	19	95.0	2226	1	POLG_HPAV8	P26582	hepatitis a
547	19	95.0	920	1	DMF1_SCHPO	P78953	schizosacch	620	19	95.0	2227	1	POLG_HPAV8	P06617	hepatitis a
548	19	95.0	921	1	PMW1_NEUCR	P707038	neurospora	621	19	95.0	2227	1	POLG_HPAV8	P06441	hepatitis a
549	19	95.0	922	1	SVI_BACSU	O45477	bacillus eu	622	19	95.0	2230	1	POLG_HPAV8	P14553	simian hepa
550	19	95.0	921	1	DPO1_RICFE	O9raa9	rickettsia	623	19	95.0	2314	1	PTP2_HUMAN	P23471	homo sapien
551	19	95.0	929	1	YDM6_SCHPO	O87137	schizosacch	624	19	95.0	2316	1	PTP2_RAT	Q62656	rattus norv
552	19	95.0	937	1	MSH2_ARATH	O24617	arabidopsis	625	19	95.0	2649	1	P285_HUMAN	Q9byk8	homo sapien
553	19	95.0	937	1	SVL_METH	O27552	methanobact	626	19	95.0	2725	1	FLNC_HUMAN	Q14115	homo sapien
554	19	95.0	942	1	MSH2_MAIZE	O9xgc9	zea mays (m	627	19	95.0	2774	1	MAPA_RAT	P34926	rattus norv
555	19	95.0	952	1	LYAG_HUMAN	P10253	homo sapien	628	19	95.0	2805	1	MAPA_HUMAN	P78559	homo sapien
556	19	95.0	953	1	SVV_VIBCH	O9kp73	vibrio chol	629	19	95.0	3412	1	POLG_TBVEV	P07720	t genome po
557	19	95.0	966	1	VIL3_ARATH	O81645	arabidopsis	630	19	95.0	3414	1	POLG_LANVT	P29837	t genome po
558	19	95.0	969	1	SVL_MYCTU	P17698	mycobacteri	631	19	95.0	3414	1	POLG_TBVEV	Q01299	t genome po
559	19	95.0	971	1	ANPN_HASCO	O10737	haemochus	632	19	95.0	3414	1	POLG_TBVEV	P14336	t genome po
560	19	95.0	972	1	MTRA_YEAST	Q99189	saccharomyc	633	19	95.0	3415	1	POLG_POWVL	Q04538	t genome po
561	19	95.0	973	1	UDV3_RHILU	Q98m36	rhizobium 1	634	19	95.0	3803	1	TRAI_DROME	Q8i8u7	drosophila
562	19	95.0	974	1	PODJ_CAUCR	O9z988	caulobacter	635	19	95.0	3859	1	TRAP_HUMAN	Q9y4a5	homo sapien
563	19	95.0	976	1	VIL2_ARATH	O81644	arabidopsis	636	19	95.0	4447	1	PKSK_BACSU	P40803	bacillus su
564	19	95.0	980	1	POLG_LIV	P22338	louping ill	637	19	95.0	5179	1	MUC2_HUMAN	Q02817	homo sapien
565	19	95.0	987	1	UVRA_XANCP	O8pa26	xanthomonas	638	19	95.0	5217	1	HTS1_COCCA	Q01886	cochliobolu
566	19	95.0	988	1	UVRA_XANCP	O8pbh3	xanthomonas	639	19	95.0	5262	1	MLL2_HUMAN	Q14686	homo sapien
567	19	95.0	992	1	UVRA_MICLU	P13567	micrococcus	640	19	95.0	5596	1	MDN1_HUMAN	Q9nu22	homo sapien
568	19	95.0	997	1	CBF_MOUSE	P53569	mus musculus	641	19	95.0	5596	1	CALM_TETH	Q05055	tetrahymena
569	19	95.0	998	1	CBF_HUMAN	O03701	homo sapien	642	18	90.0	58	1	YOZD_BACSU	Q31863	bacillus su
570	19	95.0	1012	1	DPOG_PICPA	O01941	pichia past	643	18	90.0	71	1	EX7S_STRA3	Q8e6m0	streptococc
571	19	95.0	1021	1	MAPA_PICPA	O9gyr6	mus musculus	644	18	90.0	79	1	DLTC_STRA3	Q8vm65	streptococc
572	19	95.0	1063	1	PRL1_YEAST	P12383	saccharomyc	645	18	90.0	79	1	EX7S_LACLA	Q9ch83	lactococcus
573	19	95.0	1065	1	SNCS_SCHPO	O13110	schizosacch	646	18	90.0	81	1	YKJA_ECOLI	P76036	escherichia
574	19	95.0	1086	1	RNC_CAEEL	O01326	caenorhabdi	647	18	90.0	88	1	RS15_PHOLU	P41120	photorhabdu
575	19	95.0	1095	1	AT9B_HUMAN	O43861	homo sapien	648	18	90.0	92	1	YFCL_ECOLI	P76496	escherichia
576	19	95.0	1095	1	AT9B_MOUSE	P98195	mus musculus	649	18	90.0	94	1	SELK_HUMAN	Q9y6d0	homo sapien
577	19	95.0	1117	1	CVT4_NEUCR	P47950	neurospora	650	18	90.0	94	1	SELK_MOUSE	Q9j1j1	mus musculus
578	19	95.0	1123	1	V120_HSV11	P10221	herpes simp	651	18	90.0	94	1	SELK_RAT	P59798	rattus norv
579	19	95.0	1125	1	MAP4_MOUSE	P27546	mus musculus	652	18	90.0	95	1	RS6_BACHD	Q9k5n8	bacillus ha
580	19	95.0	1144	1	FLNC_MOUSE	O8vbx6	mus musculus	653	18	90.0	98	1	YCII_ECOLI	P31070	escherichia
581	19	95.0	1146	1	YHC3_YEAST	P38742	saccharomyc	654	18	90.0	101	1	CYAV_HAEIN	P71358	haemophilus
582	19	95.0	1155	1	IFP2_METJA	O57710	methanococc	655	18	90.0	108	1	PRIB_NITEU	O82xq7	nitrosomona
583	19	95.0	1167	1	CAGA_HELPJ	O9zlt1	helicobacte	656	18	90.0	110	1	VAG1_TOBAC	O82702	nicotiana t
584	19	95.0	1173	1	UB4B_MOUSE	O9e00	mus musculus	657	18	90.0	111	1	VAG2_TOBAC	O82703	nicotiana t
585	19	95.0	1200	1	MOG5_CAEEL	O09530	caenorhabdi	658	18	90.0	117	1	RNPA_LACLA	Q9c373	lactococcus
586	19	95.0	1211	1	DP3A_HELPJ	O9zlf9	helicobacte	659	18	90.0	133	1	S3AD_BACSU	P49781	bacillus su
587	19	95.0	1211	1	DP3A_HELPJ	P56157	helicobacte	660	18	90.0	134	1	YK49_STRCO	P16251	streptomyce
588	19	95.0	1239	1	TOP2_CRIFA	P27570	crithidia f	661	18	90.0	136	1	CDD_BACPY	Q9s3m0	bacillus ps
589	19	95.0	1302	1	UB4B_HUMAN	O95155	homo sapien	662	18	90.0	136	1	RL19_XYLFA	Q9p326	xylella fas
590	19	95.0	1311	1	SPR4_CAEEL	O17582	caenorhabdi	663	18	90.0	136	1	RL19_XYLFA	Q87f53	xylella fas
591	19	95.0	1331	1	CVAB_LEIDO	O25463	leishmania	664	18	90.0	141	1	ALL2_TYRPU	O02380	tyrophagus
592	19	95.0	1355	1	DP3A_SYNY3	P74750	synchocyst	665	18	90.0	141	1	RISB_METJA	Q57751	methanococc
593	19	95.0	1371	1	VCAP_HSVSA	O00999	herpesvirus	666	18	90.0	144	1	Y991_METJA	Q58398	methanococc
594	19	95.0	1423	1	ALBU_PETMA	Q91274	petromyzon	667	18	90.0	145	1	YK14_VIBVU	Q8db14	vibrio vuln
595	19	95.0	1426	1	NPH4_HUMAN	O75161	homo sapien	668	18	90.0	147	1	CALM_KLUULA	O60041	kluyveromyc
596	19	95.0	1444	1	DPO3_LISIN	Q92634	listeria in	669	18	90.0	147	1	CALM_YEAST	P06787	saccharomyc
597	19	95.0	1450	1	DPO3_LISMO	O8y7g1	listeria mo	670	18	90.0	147	1	YESE_BACSU	O31511	bacillus su
598	19	95.0	1477	1	YORI_YEAST	P53049	saccharomyc	671	18	90.0	148	1	CAL2_PETHY	P27163	petunia hyb
599	19	95.0	1515	1	YCFI_YEAST	P39109	saccharomyc	672	18	90.0	148	1	CAL2_DROME	P49258	drosophila
600	19	95.0	1522	1	MRP3_RAT	O88563	rattus norv	673	18	90.0	148	1	CALM_ACHKL	P15094	achlya kleb
601	19	95.0	1524	1	Y133_HUMAN	Q14146	homo sapien	674	18	90.0	148	1	CALM_BLAEM	Q9hfy6	blastoclad
602	19	95.0	1527	1	MRP3_HUMAN	O15438	homo sapien	675	18	90.0	148	1	CALM_DROME	P07181	drosophila
603	19	95.0	1535	1	LML1_CAEEL	P18823	caenorhabdi	676	18	90.0	148	1	CALM_EUEEL	P02594	electrophor
604	19	95.0	1538	1	LHR_ECOLI	Q30015	escherichia	677	18	90.0	148	1	CALM_EUGRA	P11118	eugenia gra
605	19	95.0	1609	1	CTP1_MYLE	O53114	mycobacteri	678	18	90.0	148	1	CALM_EUGGR	P13565	hordeum vul
606	19	95.0	1679	1	GCPI_MOUSE	O8cbg3	mus musculus	679	18	90.0	148	1	CALM_HORVU	P02593	homo sapien
607	19	95.0	1906	1	YFAO_ANASP	O8vm40	anabaena sp	680	18	90.0	148	1	CALM_HUMAN	P27161	lycopersico
608	19	95.0	1961	1	MYH9_RAT	Q62812	rattus norv	681	18	90.0	148	1	CALM_LYCES	P41040	zea mays (m
609	19	95.0	2027	1	DOC3_MOUSE	Q8c1q7	mus musculus	682	18	90.0	148	1	CALM_MAIZE	P02596	metridium s
610	19	95.0	2030	1	DOC3_HUMAN	Q8i4d9	homo sapien	683	18	90.0	148	1	CALM_METSE	P07463	paramecium
611	19	95.0	2059	1	TEGU_HSV7J	P52362	human herpe	684	18	90.0	148	1	CALM_ORYSA	P02595	phatopect
612	19	95.0	2061	1	MYOF_HUMAN	O9nm21	homo sapien	685	18	90.0	148	1	CALM_PATSP	P27165	phytophthor
613	19	95.0	2165	1	RPL1_HRSVA	P28897	human respi	686	18	90.0	148	1	CALM_PHTYN	P11121	pyridae sp
614	19	95.0	2208	1	POLN_MANCV	Q69014	manchester	687	18	90.0	148	1	CALM_PYUSP	P13868	solanum tub
615	19	95.0	2210	1	RPOE_TACV	P20430	tacaribe vi	688	18	90.0	148	1	CALM_SOLTU	P04353	spinacia ol
616	19	95.0	2222	1	DPOE_YEAST	P21951	saccharomyc	689	18	90.0	148	1	CALM_SPTOL	P21251	stichopus j
617	19	95.0	2226	1	POLG_HPAV2	P26580	hepatitis a	690	18	90.0	148	1	CALM_STIJA		

691	18	90.0	148	1	CALM_STYLE	P27166 stylonychia	764	18	90.0	229	1	RADB_THEAC	Q9hjd3 thermoplasm
692	18	90.0	148	1	CALM_TETPY	P02598 tetrahymena	765	18	90.0	230	1	ISPD_SYNV3	P74323 synechocyst
693	18	90.0	148	1	CALM_CHICK	P03597 gallus gall	766	18	90.0	231	1	FP_CRIMI	P15697 cricetus
694	18	90.0	149	1	CALM_CANAL	P23286 candida alb	767	18	90.0	234	1	SAMP_MESAU	P07629 mesocricetu
695	18	90.0	149	1	CALM_WHEAT	P04464 triticum ae	768	18	90.0	235	1	ICLN_CANFA	P35521 canis famli
696	18	90.0	149	1	MLE3_HUMAN	P06741 homo sapien	769	18	90.0	236	1	PUR7_HELHP	Q7viq6 helicobacte
697	18	90.0	149	1	MLE3_RABIT	P02603 cryptolagus	770	18	90.0	237	1	PUR7_METAC	Q8t189 methanosarc
698	18	90.0	149	1	MLE3_RAT	P02601 rattus norv	771	18	90.0	237	1	PUR7_METWA	Q8pyk6 methanosarc
699	18	90.0	150	1	CALM_SGPO	P05933 schizosacch	772	18	90.0	238	1	RL1_ANASP	Q8ylj7 anabaena sp
700	18	90.0	150	1	V351_METJA	P05797 methanococc	773	18	90.0	238	1	Y538_CHLTR	O84543 chlamydia t
701	18	90.0	151	1	CALM_DICDI	P02599 dictyosteli	774	18	90.0	238	1	Y825_CHLAMU	Q8pjka chlamydia m
702	18	90.0	151	1	CP2B_DROME	P13486 drosophila	775	18	90.0	240	1	PUR7_COXBU	Q83ca8 coxiella bu
703	18	90.0	151	1	Y4G6_RHIME	P13486 rhizobium m	776	18	90.0	242	1	MTGA_ECOLI	P46022 escherichia
704	18	90.0	152	1	AAC6_ENTAE	P50858 enterobacte	777	18	90.0	242	1	MTGA_KLEOX	Q48465 klebsiella
705	18	90.0	152	1	CALF_NAAGR	P53440 naegleria g	778	18	90.0	242	1	NPDI_THETN	Q8rgn6 thermoanaer
706	18	90.0	155	1	Y463_FUSNR	Q8X955 fusobacteri	779	18	90.0	242	1	PDJX_NEIMA	Q8rgv9 neisseria m
707	18	90.0	159	1	TPCS_HUMAN	P03585 homo sapien	780	18	90.0	242	1	RRPO_SCVLA	P23172 saccharomyc
708	18	90.0	159	1	TPCS_MOUSE	P20801 mus musculu	781	18	90.0	244	1	TRCB_HUMAN	Q89426 homo sapien
709	18	90.0	159	1	TPCS_PIG	P02587 sus scrofa	782	18	90.0	244	1	YLS2_ARCTU	O28148 archaeoglob
710	18	90.0	159	1	TPCS_RABIT	P02586 cryptolagus	783	18	90.0	245	1	YCDX_ECOLI	P75914 escherichia
711	18	90.0	160	1	TPCS_ANGAN	P81660 anguilla an	784	18	90.0	246	1	AQPM_METH	O26206 methanobact
712	18	90.0	162	1	AROK_LACLA	Q9ceul lactococcus	785	18	90.0	246	1	AQPM_METTM	Q9C425 methanobact
713	18	90.0	162	1	CALM_CHLRE	P04352 chlamydomon	786	18	90.0	246	1	HEM4_CHLVI	Q59335 chlorobium
714	18	90.0	162	1	TPCS_CHICK	P03588 gallus gall	787	18	90.0	246	1	MTGA_HAETN	P44890 haemophilus
715	18	90.0	162	1	TPCS_MELGA	P10246 meleagris g	788	18	90.0	246	1	TRPC_SULTO	Q972a1 sulfolobus
716	18	90.0	162	1	TPCS_RANES	P03589 rana esculu	789	18	90.0	248	1	Y4EL_RHISN	P55435 rhizobium s
717	18	90.0	166	1	Y551_SULSO	Q9uwv6 sulfolobus	790	18	90.0	251	1	TPIS_PSESM	Q87wq1 pseudomonas
718	18	90.0	166	1	Y628_METJA	Q58045 methanococc	791	18	90.0	251	1	TPIS_PSESY	P95576 pseudomonas
719	18	90.0	169	1	IPVR_ANASP	P80562 anabaena sp	792	18	90.0	252	1	RECO_CHLPN	Q927w5 chlamydia p
720	18	90.0	169	1	IPVR_SYNV3	P80507 synechocyst	793	18	90.0	252	1	TRPC_THENA	Q56319 thermotoga
721	18	90.0	171	1	NSG2_HUMAN	P94328 homo sapien	794	18	90.0	253	1	YTBQ_BACSU	P33560 bacillus su
722	18	90.0	171	1	NSG2_MOUSE	P47759 mus musculu	795	18	90.0	254	1	AROD_NEIMB	Q9iyt0 neisseria m
723	18	90.0	171	1	YBP2_ACIAM	P32986 acidianus a	796	18	90.0	254	1	PMW_YEAST	P07283 saccharomyc
724	18	90.0	172	1	IPVR_SYNEL	P50157 ambystoma m	797	18	90.0	257	1	RPSE_BACAA	Q8K5j6 bacillus an
725	18	90.0	173	1	TCPA_ANEME	P40423 drosophila	798	18	90.0	257	1	UXUR_ECOLI	P39161 escherichia
726	18	90.0	174	1	MLRN_DROME	P50157 ambystoma m	799	18	90.0	258	1	EXUR_ECOLI	P42608 escherichia
727	18	90.0	177	1	IPVR_HALN1	Q9hef3 halobacteri	800	18	90.0	259	1	EXUR_ERWCH	Q9x9e0 erwinia chr
728	18	90.0	181	1	IPVR_THEAQ	P96078 thermus aqu	801	18	90.0	260	1	RS3_ANASP	Q8yp15 anabaena sp
729	18	90.0	184	1	Y803_CHLPN	P927a3 chlamydia p	802	18	90.0	260	1	TRPC_BACTN	Q8aad6 bacteroides
730	18	90.0	185	1	CBX1_HUMAN	P23197 homo sapien	803	18	90.0	261	1	RPSE_BACSU	P06574 bacillus su
731	18	90.0	185	1	DPI_HUMAN	Q00765 homo sapien	804	18	90.0	264	1	YNI6_YEAST	P33573 saccharomyc
732	18	90.0	185	1	DPI_MOUSE	Q00870 mus musculu	805	18	90.0	265	1	GIR2_YEAST	P33768 saccharomyc
733	18	90.0	185	1	NSGI_HUMAN	P42857 homo sapien	806	18	90.0	266	1	UXUR_HAETN	P44487 haemophilus
734	18	90.0	185	1	NSGI_MOUSE	Q62092 mus musculu	807	18	90.0	267	1	APHC_HUMAN	Q9nun7 homo sapien
735	18	90.0	185	1	RFBC_ECOLI	P37745 escherichia	808	18	90.0	267	1	RL4_SULSO	Q8uxa6 sulfolobus
736	18	90.0	185	1	Y56A_THEMA	P58008 thermotoga	809	18	90.0	267	1	YG95_HAETN	Q48215 haemophilus
737	18	90.0	189	1	KTHV_THEAC	Q9hlz2 thermoplasm	810	18	90.0	268	1	HIS9_BACSU	O34411 bacillus su
738	18	90.0	197	1	Y249_METH	O26351 methanobact	811	18	90.0	268	1	THIM_ARCFU	Q38204 archaeoglob
739	18	90.0	199	1	TDX2_BRUMA	Q17172 brugia mala	812	18	90.0	269	1	TRPA_BACST	P39867 bacillus st
740	18	90.0	200	1	ESTE_VIBMI	Q07792 vibrio mimi	813	18	90.0	270	1	RPNC_SCHFO	P50524 schizosacch
741	18	90.0	201	1	RETE_MOUSE	Q00724 mus musculu	814	18	90.0	270	1	YCB7_METUA	Q59087 acinetobact
742	18	90.0	201	1	RETE_RAT	P04916 rattus norv	815	18	90.0	272	1	3DHQ_ACICA	Q50501 rana esculu
743	18	90.0	202	1	COAE_CHLTR	O84499 chlamydia t	816	18	90.0	272	1	AQPA_RANES	O14415 candida alb
744	18	90.0	202	1	HIS5_METAC	Q8ce91 methanosarc	817	18	90.0	275	1	SC65_CANAL	Q30083 archaeoglob
745	18	90.0	202	1	HIS5_METMA	Q8pvd5 methanosarc	818	18	90.0	277	1	Y154_ARCFU	Q30083 archaeoglob
746	18	90.0	202	1	YC55_MYCTU	Q11063 mycobacteri	819	18	90.0	278	1	P128_ARATH	Q8zvx8 arabidopsis
747	18	90.0	206	1	CNC2_MOUSE	Q9db76 mus musculu	820	18	90.0	279	1	YD15_STAAM	Q8nrw0 staphylococ
748	18	90.0	208	1	CNC2_HUMAN	Q9v3b6 homo sapien	821	18	90.0	279	1	YE25_STAAM	Q8nu62 staphylococ
749	18	90.0	208	1	RL1_PEA	P43208 pisum sativ	822	18	90.0	280	1	P127_ARATH	P33004 arabidopsis
750	18	90.0	208	1	RR3_GRATE	P16631 gracilaria	823	18	90.0	281	1	GLPF_ECOLI	P11244 escherichia
751	18	90.0	209	1	TRPE_METKA	Q8cx29 methanopyru	824	18	90.0	281	1	GLPF_SHIFL	P31140 shigella fl
752	18	90.0	216	1	DEF1_BIFLO	Q8G534 bifidobacte	825	18	90.0	282	1	P1P1_ATRCA	P42767 atrilex ca
753	18	90.0	217	1	RECA_ARTAU	Q9xej6 arthrobacte	826	18	90.0	283	1	LGT_HELPU	P44420 haemophilus
754	18	90.0	221	1	MLR_DROME	P18432 drosophila	827	18	90.0	283	1	TVSY_HAETN	P57808 pasteurella
755	18	90.0	223	1	DEOC_MYCGE	P47296 mycoplasma	828	18	90.0	283	1	TVSY_PASMU	O25609 helicobacte
756	18	90.0	223	1	SAMP_CAVPO	P49255 cavia porce	829	18	90.0	284	1	LGT_HELPY	Q8tvh1 methanopyru
757	18	90.0	223	1	SAMP_HUMAN	P02743 homo sapien	830	18	90.0	284	1	NAD6_METKA	P59594 sulfolobus
758	18	90.0	224	1	DEOC_MYCPN	P03924 mycoplasma	831	18	90.0	284	1	Y066_SULSO	P43287 arabidopsis
759	18	90.0	224	1	SAMP_PIG	O19063 sus scrofa	832	18	90.0	285	1	P122_ARATH	P30302 arabidopsis
760	18	90.0	226	1	ORGO_SALTY	P58654 salmonella	833	18	90.0	285	1	P123_ARATH	P3285 arabidopsis
761	18	90.0	228	1	ISPD_ANASP	Q8vlyx9 anabaena sp	834	18	90.0	286	1	P111_ARATH	Q06611 arabidopsis
762	18	90.0	228	1	IPVR_METH	O26232 methanobact	835	18	90.0	286	1	P112_ARATH	Q08733 arabidopsis
763	18	90.0	228	1	SAMP_RAT	P23680 rattus norv	836	18	90.0	286	1	P113_ARATH	

837	18	90.0	286	1	PI25 ARATH	Q9sv31 arabidopsis	910	18	90.0	363	1	LEU3_BUCAP	O85064 buchnera ap
838	18	90.0	286	1	PI11 LYCES	Q08451 lycopersico	911	18	90.0	363	1	MURG_ENTFA	O07109 enterococu
839	18	90.0	287	1	PI14 ARATH	Q39196 arabidopsis	912	18	90.0	364	1	AAT_PVRKO	Q93744 pyrococcus
840	18	90.0	287	1	PI15 ARATH	Q81aa6 arabidopsis	913	18	90.0	364	1	YM28 MYCTU	Q10512 mycobacteri
841	18	90.0	287	1	PI21 ARATH	Q43286 arabidopsis	914	18	90.0	365	1	HI82 BORBR	Q7wdy3 bordetella
842	18	90.0	289	1	PI26 ARATH	Q9zv07 arabidopsis	915	18	90.0	365	1	HI82 BORPA	Q7w2y3 bordetella
843	18	90.0	289	1	PI26 ARATH	Q9zv07 arabidopsis	915	18	90.0	365	1	HI82 BORPA	Q7w2y3 bordetella
844	18	90.0	289	1	PI26 ARATH	Q9zv07 arabidopsis	915	18	90.0	365	1	HI82 BORPA	Q7w2y3 bordetella
845	18	90.0	291	1	EGC1 LISMO	P25794 pisum sativ	916	18	90.0	365	1	HI82 BORPE	P48293 spirulina p
846	18	90.0	291	1	PI24 ARATH	Q8y680 listeria mo	917	18	90.0	365	1	RECA SPIPL	P50213 bos taurus
847	18	90.0	294	1	ISPA_BUCAP	Q9f553 arabidopsis	918	18	90.0	366	1	IDHA BOVIN	P41563 hmo sapien
848	18	90.0	294	1	ISPA_BUCAP	Q8k9a0 buchnera ap	919	18	90.0	366	1	IDHA HUMAN	P48293 spirulina p
849	18	90.0	296	1	NI11 ARATH	Q8vz01 arabidopsis	920	18	90.0	369	1	METB_HAEIN	P44502 haemophilus
850	18	90.0	296	1	NI11 ARATH	Q8vz01 arabidopsis	920	18	90.0	369	1	METB_HAEIN	P44502 haemophilus
851	18	90.0	298	1	YK23 ARCFU	O82556 archaeoglob	921	18	90.0	369	1	RECA CORPS	P48288 corynebacte
852	18	90.0	299	1	PSD_CHLCV	Q82113 chlamydophi	922	18	90.0	369	1	VP6_AHSV3	Q64909 african hor
853	18	90.0	300	1	Y025 FOPVP	O10298 orgyia pseu	923	18	90.0	369	1	VP6_AHSV6	Q64913 african hor
854	18	90.0	302	1	RP04 FOPVP	Q9j540 fowlpox vir	924	18	90.0	371	1	Y1B0 CIOAB	Q04354 clostridium
855	18	90.0	302	1	EST_ACIUN	P18773 acinetobact	925	18	90.0	372	1	BIOF_METJA	Q58694 methanococ
856	18	90.0	303	1	T2RD HUMAN	Q9nyv9 homo sapien	926	18	90.0	372	1	DEGT_BACST	P15263 bacillus st
857	18	90.0	303	1	T2RD HUMAN	Q9nyv9 homo sapien	926	18	90.0	372	1	DEGT_BACST	P15263 bacillus st
858	18	90.0	303	1	T2RD HUMAN	Q9nyv9 homo sapien	926	18	90.0	372	1	DEGT_BACST	P15263 bacillus st
859	18	90.0	304	1	DCAS_AGRSTU	Q44185 agrobacteri	927	18	90.0	376	1	RECA CORGL	P42442 corynebacte
860	18	90.0	304	1	DCAS_AGRSTU	Q44185 agrobacteri	927	18	90.0	376	1	RECA CORGL	P42442 corynebacte
861	18	90.0	304	1	DCAS_AGRSTU	Q44185 agrobacteri	927	18	90.0	376	1	RECA CORGL	P42442 corynebacte
862	18	90.0	304	1	DCAS_AGRSTU	Q44185 agrobacteri	927	18	90.0	376	1	RECA CORGL	P42442 corynebacte
863	18	90.0	304	1	DCAS_AGRSTU	Q44185 agrobacteri	927	18	90.0	376	1	RECA CORGL	P42442 corynebacte
864	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
865	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
866	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
867	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
868	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
869	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
870	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
871	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
872	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
873	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
874	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
875	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
876	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
877	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
878	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
879	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
880	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
881	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
882	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
883	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
884	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
885	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
886	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
887	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
888	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
889	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
890	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
891	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
892	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
893	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
894	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
895	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
896	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
897	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
898	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
899	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
900	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
901	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
902	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
903	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
904	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
905	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
906	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
907	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
908	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus
909	18	90.0	305	1	TR2D RAT	Q8nwv8 staphylococ	928	18	90.0	380	1	YH07_PASMU	Q9ra89 lactococcus

```

983 18 90.0 474 1 SYC PYRAE
984 18 90.0 477 1 GUNJ ECOLI
985 18 90.0 484 1 GUNA XANCP
986 18 90.0 490 1 DCLY BACSU
987 18 90.0 498 1 SECY MYCGA
988 18 90.0 499 1 SYPA THEAC
989 18 90.0 500 1 GALT CAEEL
990 18 90.0 500 1 NU4C ARATH
991 18 90.0 506 1 PUR9 ANRSP
992 18 90.0 507 1 C4DE DROME
993 18 90.0 509 1 ATPA RHOC
994 18 90.0 511 1 YE08 YEAST
995 18 90.0 513 1 Y538 RICPR
996 18 90.0 529 1 PUR9 CAUCR
997 18 90.0 529 1 VSM6 TRYBB
998 18 90.0 529 1 YB89 YEAST
999 18 90.0 532 1 FM03 CAVPO
1000 18 90.0 536 1 ENTE ECO57

Q8xyd8 pyrobaculum
P42590 escherichia
P19487 xanthomonas
P21885 bacillus su
O52351 mycoplasma
P57693 thermoplas
Q956t2 caenorhabdi
P28288 arabidopsis
Q8yej2 anabaena sp
O46051 drosophila
P72245 rhodobacter
P40051 saccharomyc
Q9zd12 rickettsia
Q9aby4 c bifunctio
P26334 trypanosoma
P38140 saccharomyc
P49109 cavia porce
Q8xbv3 escherichia

```

## ALIGNMENTS

```

RESULT 1
NS2_MYCTU STANDARD; PRT; 19 AA.
NC P81136;
YT 15-JUL-1999 (Rel. 38, Created)
YT 15-JUL-1999 (Rel. 38, Last sequence update)
YT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30 kDa non-secretory protein 2 (fragment).
NS Mycobacterium tuberculosis.
XC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
XC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
DX NCBI_TaxID=1773;
UN [1]
CP SEQUENCE.
IC STRAIN=H37Rv;
LA Prasad H.K., Annapurna P.S.;
L Submitted (DEC-1997) to Swiss-Prot.
C -I- SIMILARITY: STRONG, TO THE N-TERMINALS OF C.ELEGANS F20A1.4
C AND H.INFLUENZA H10967.
C -I- CAUTION: We are unable to find this protein in the translation of
C the genome of strain H37Rv.
T NON_TER 19
Q SEQUENCE 19 AA; 2211 MW; ABC1854BF1FF1F70 CRC64;

```

```

Query Match 95.0%; Score 19; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Y 2 VAEF 5
b 11 VAEF 14

```

```

RESULT 2
ECF_TREPA STANDARD; PRT; 59 AA.
C O83263;
T 15-DEC-1998 (Rel. 37, Created)
T 15-DEC-1998 (Rel. 37, Last sequence update)
T 16-OCT-2001 (Rel. 40, Last annotation update)
E Preprotein translocase secE subunit.
N SECE OR TP0235.
S Treponema pallidum.
C Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
X NCBI_TaxID=160;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=Nichols;
X MEDLINE=98332770; PubMed=965876;
A Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,

```

```

RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.,
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.",
RL Science 281:375-388(1998).
CC -I- FUNCTION: Essential for protein export.
CC -I- SUBCELLULAR LOCATION: Tail-anchored membrane protein (Potential).
CC -I- SIMILARITY: Belongs to the secE/SEC61-gamma family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE001205; AAC65223.1; -.
CC PIR; E71349; E71349.
CC TIGR; TP0235; -.
CC InterPro; IPR001901; SecE.
CC InterPro; IPR005807; SecE_bac.
CC Pfam; PF00584; SecE; 1.
CC TIGRFAMs; TIGR00964; 3a0501s06; 1.
CC PROSITE; PS01667; SECE_SEC61G; 1.
KW Protein transport; Translocation; Transmembrane; Complete proteome.
FT TRANSMEM 39 59 POTENTIAL.
SQ SEQUENCE 59 AA; 6789 MW; 9AC35BA8F48B2A7F CRC64;

```

```

Query Match 95.0%; Score 19; DB 1; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 VAEF 5
Db 12 VAEF 15

```

```

RESULT 3
YA90_ARCFU STANDARD; PRT; 59 AA.
ID YA90_ARCFU
AC O29175;
AT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hypothetical UPF0165 protein AF1090.
GN AF1090.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty E.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.",
RL Nature 390:364-370(1997).

```

-!- SIMILARITY: Belongs to the UPF0165 family.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; AE001028; AAB90157.1; -.  
 PIR; A69386; A69386.

TIGR; AF1090; -.

InterPro; IPR008203; DUF104.

InterPro; IPR008204; DUF104\_N.

Pfam; PF01954; DUF104; 1.

ProDom; PD005964; DUF104\_N; 1.

Hypothetical protein; Complete proteome.

SEQUENCE 59 AA; 6887 MW; B72E847382B56B6 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 59;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5

|||||

51 VAEF 54

#### SUIT 4

51 HAEIN

Y451 HAEIN STANDARD; PRT; 63 AA.

P43998;

01-NOV-1995 (Rel. 32, Created)

01-NOV-1995 (Rel. 32, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

Hypothetical protein HI0451.

HI0451.

Haemophilus influenzae.

Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

Pasteurellaceae; Haemophilus.

NCBI\_TaxID=727;

[1]

SEQUENCE FROM N.A.

STRAIN=Rd / KW20 / ATCC 51907;

MEDLINE=95350630; PubMed=7542800;

Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,

Fine L.D., Pritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,

Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

Venter J.C.;

"Whole-genome random sequencing and assembly of Haemophilus influenzae

Rd.";

Science 269:496-512(1995).

-----

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; U32728; AAC22109.1; -.

PIR; G64007; G64007.

TIGR; HI0451; -.

Hypothetical protein; Complete proteome.

SEQUENCE 63 AA; 7047 MW; 36814B60AC306B67 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 63;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5

|||||

Db 27 VAEF 30

#### RESULT 5

RS15 HAEIN

ID RS15 HAEIN STANDARD; PRT; 88 AA.

AC P44393;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 30S ribosomal protein S15.

GN (RPSO-A OR RPS15-A OR H11328) AND (RPSO-B OR RPS15-B OR H11468).

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Haemophilus.

OK NCBI\_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Rd / KW20 / ATCC 51907;

RC MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Gocayne J.D.,

McKenney K., Sutton G., Fitzhugh W., Fields C.A., Kelliey J.M.,

Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,

Fine L.D., Pritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,

Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

Venter J.C.;

"Whole-genome random sequencing and assembly of Haemophilus influenzae

Rd.";

Science 269:496-512(1995).

-!- FUNCTION: This protein is one of the 16S ribosomal RNA binding

proteins (By similarity).

-!- SIMILARITY: Belongs to the S15P family of ribosomal proteins.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; U32812; AAC22973.1; -.

EMBL; U32825; AAC23117.1; -.

PIR; H64116; H64116.

HSPP; P05766; 1A32.

TIGR; H11328; -.

TIGR; H11468; -.

InterPro; IPR00589; Ribosomal\_S15.

InterPro; IPR005290; Ribosomal\_S15\_b.

Pfam; PF00312; Ribosomal\_S15; 1.

ProDom; PD157043; RS15\_bact; 1.

TIGRPMs; TIGR00952; S15\_bact; 1.

PROSITE; PS00362; RIBOSOMAL\_S15; 1.

Ribosomal protein; rRNA-binding; Complete proteome.

INIT\_MET 0 0 BY SIMILARITY.

SEQUENCE 88 AA; 10064 MW; 563BAD2B8B9A7043 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 88;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5

|||||

Db 11 VAEF 14

```

OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Karvase A.R., Dougherty B.A., Tomb J.-P., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodok A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii".
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: BELONGS TO THE UPF0166 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67593; AAB99550.1;
DR PIR; C64490; C64490.
DR TIGR; MJ1524; -.
DR InterPro; IPR003793; DUF190.
DR Pfam; PF02641; DUF190; 1.
RW Hypothetical protein; Complete proteome.
SQ SEQUENCE 108 AA; 12398 MW; D3E286AFB9D4AF8B CRC64;

Query Match 95.0%; Score 19; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 52 VAEF 55

RESULT 8
KEDD ACTSL STANDARD; PRT; 114 AA.
AC P41249;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apokedarcidin.
OS Actinomycete sp. (strain L585-6 / ATCC 53650).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales.
OX NCBI_TaxID=38989;
RN [1]
RP SEQUENCE.
RC STRAIN=ATCC 53650 / L585-6;
RX Hofstead S.J., Matson J.A., Malacko A.R., Marguardt H.;
RT "Kedarcidin, a new chromoprotein antitumor antibiotic. II. Isolation,
RT purification and physico-chemical properties.";
RL J. Antibiot. 45:1250-1254(1992).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93376732; PubMed=8367457;
RA Zein N., Casazza A.M., Doyle T.W., Leet J.E., Schoeder D.R.,
RA Solomon W., Nadler S.G.;
RT "Selective proteolytic activity of the antitumor agent kedarcidin.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:8009-8012(1993).
RN [3]

```

STRUCTURE BY NMR.  
 STRAIN=ATCC 53650 / L585-6;  
 MEDLINE=95001848; PubMed=7918358;  
 Constantine K.L., Colson K.L., Wittekind M., Friedrichs M.S.,  
 Zein N., Tuttle J., Langley D.R., Leet J.E., Schroeder D.R., Lam K.S.,  
 Farmer B.T. II, Metzler W.J., Brucoleri R.E., Mueller L.,  
 "Sequential 1H, 13C, and 15N NMR assignments and solution  
 conformation of apokadarcidin.";  
 Biochemistry 33:11438-11452(1994).  
 -!- FUNCTION: BINDS NON-COVALENTLY TO AN ENEDIYNE CHROMOPHORE WHICH IS  
 THE CYTOTOXIC AND MUTAGENIC COMPONENT OF THE ANTIBIOTIC. THE  
 CHROMOPHORE CLEAVES DUPLEX DNA SITE-SPECIFICALLY IN A SINGLE-  
 STRANDED MANNER. THE APOPROTEIN CLEAVES PROTEINS SELECTIVELY, IN  
 PARTICULAR HIGHLY BASIC HISTONES, WITH H1 PROTEINS BEING CLEAVED  
 THE MORE READILY.  
 -!- DOMAIN: THIS PROTEIN CONSISTS OF AN IMMUNOGLOBULIN-LIKE SEVEN-  
 STRANDED ANTIPARALLEL BETA-BARREL DOMAIN LINKED TO A SUBDOMAIN  
 COMPOSED OF TWO BETA-HAIRPIN RIBBONS.  
 -!- SIMILARITY: Belongs to the neocarzinostatin family.  
 PDB; JAKP; 3I-AUG-94.  
 InterPro; IPR002186; Neocarzinostat.  
 Pfam; PF00960; Neocarzinostat; 1.  
 ProDom; PPO12709; Neocarzinostat; 1.  
 Antibiotic; DNA-binding; 3D-structure.  
 DISULFID 37 47  
 STRAND 4 7  
 TURN 11 13  
 STRAND 15 16  
 TURN 18 18  
 STRAND 21 24  
 STRAND 32 33  
 STRAND 36 39  
 STRAND 46 47  
 TURN 50 51  
 STRAND 55 55  
 STRAND 61 61  
 STRAND 64 64  
 STRAND 67 67  
 STRAND 71 74  
 TURN 77 79  
 STRAND 84 87  
 STRAND 94 97  
 STRAND 100 100  
 STRAND 109 110  
 STRAND 112 114  
 SEQUENCE 114 AA; 10969 MW; 190152B14E4197B4 CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 ||||  
 49 VAEF 52  
 S\_VERMO  
 INS\_VERMO STANDARD; PRT; 115 AA.  
 Q9W7R2;  
 16-OCT-2001 (Rel. 40, Created)  
 16-OCT-2001 (Rel. 40, Last sequence update)  
 10-OCT-2003 (Rel. 42, Last annotation update)  
 Insulin precursor.  
 INS.  
 Verasper moseri (Barfin flounder).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
 Pleuronectoidei; Pleuronectidae; Verasper.  
 NCBI\_TaxID=98923;  
 [1]

RP SEQUENCE FROM N.A.  
 RA Andoh T., Nagasawa H.;  
 RT "Two molecular forms of insulin from barfin flounder, Verasper moseri,  
 are derived from a single gene.";  
 RL Zool. Sci. 15:931-937(1998).  
 CC -!- FUNCTION: Insulin decreases blood glucose concentration. It  
 increases cell permeability to monosaccharides, amino acids and  
 fatty acids. It accelerates glycolysis, the pentose phosphate  
 cycle, and glycogen synthesis in liver.  
 CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two  
 disulfide bonds.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the insulin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AB029318; BAB82315.1; -;  
 DR HSSP; P01315; INPJ.  
 DR InterPro; IPR004825; Ins/IGP/relax.  
 DR Pfam; PF00049; Insulin; 1.  
 DR PRINTS; PR00277; INSULIN.  
 DR SMART; SM00078; IIGF; 1.  
 DR PROSITE; PS00262; INSULIN; 1.  
 DR Insulin family; Hormone; Glucose metabolism; Signal.  
 FT SIGNAL 1 22  
 CHAIN 23 53  
 PROPEP 56 92  
 FT CHAIN 95 115  
 FT CHAIN 32 101  
 FT DISULFID 44 114  
 FT DISULFID 100 105  
 FT SEQUENCE 115 AA; 12608 MW; 7BA2A5B568DEDBB CRC64;  
 SQ  
 Query Match 95.0%; Score 19; DB 1; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 ||||  
 79 VAEF 82  
 Db  
 RESULT 10  
 INS\_LOPPI  
 ID INS\_LOPPI STANDARD; PRT; 116 AA.  
 AC P01341;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Insulin precursor.  
 GN INS.  
 OS Lophius piscatorius (Allmouth goosefish) (Anglerfish), and  
 OS Lophius americanus (American goosefish) (Anglerfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophidae; Lophius.  
 OX NCBI\_TaxID=8074, 8073;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=L. americanus;  
 RX MEDLINE=81056434; PubMed=7001633;  
 RA Hobart P.M., Shen L.-F., Crawford R., Pictet R.L., Rutter W.J.;  
 RT "Comparison of the nucleic acid sequence of anglerfish and mammalian  
 insulin mRNA's from cloned cDNA's";  
 RL Science 210:1360-1363(1980).  
 RN [2]  
 RP SEQUENCE OF 25-54 AND 96-116.



```

C SPECIES=L.piscatorius;
X MEDLINE=70036620; PubMed=5389299;
A Neumann P.A., Koldenhof M., Humbel R.E.;
T "Amino acid sequence of insulin from the angler fish (Lophius
T borealis)".
L Hoppe-Seyler's Z. Physiol. Chem. 350:1286-1288 (1969).
C -1- FUNCTION: Insulin decreases blood glucose concentration. It
C increases cell permeability to monosaccharides, amino acids and
C fatty acids. It accelerates glycolysis, the pentose phosphate
C cycle, and glycogen synthesis in liver.
C -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
C disulfide bonds.
C -1- SUBCELLULAR LOCATION: Secreted.
C -1- SIMILARITY: Belongs to the insulin family.
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See http://www.isb-sib.ch/announce/
C or send an email to license@isb-sib.ch).
C -----
R EMBL; V00634; CAA23907.1; -.
R PIR: A01608; IPAF.
R HSPSP; P01308; IIPH.
R InterPro: IPR004825; Ins/IGF/relax.
R Pfam: PF00049; Insulin; 1.
R PRINTS; PR00277; INSULIN.
R SMART; SM00078; IIGF; 1.
R PROSITE; PS00262; INSULIN; 1.
R Insulin family; Hormone; Glucose metabolism; Signal.
T SIGNAL 1 24
T CHAIN 1 24 INSULIN B CHAIN.
T PROPEP 25 53 C PEPTIDE.
T CHAIN 56 93
T DISULFID 96 116 INSULIN A CHAIN.
T DISULFID 32 102 INTERCHAIN.
T DISULFID 44 115
T DISULFID 101 106
T SEQUENCE 116 AA; 12737 MW; C68F8F8F8183BEFE CRC64;
Y Query Match 95.0%; Score 19; DB 1; Length 116;
Best Local Similarity 100.0%; Pred.No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
b 2 VAEF 5
|
|
|
80 VAEF 83
RESULT 11
D U279 DROME STANDARD; PRT; 119 AA.
C Q9W8F3;
T 10-OCT-2003 (Rel. 42, Created)
T 10-OCT-2003 (Rel. 42, Last sequence update)
T 10-OCT-2003 (Rel. 42, Last annotation update)
E UF0279 protein CGI4505.
N CGI4505.
S Drosophila melanogaster (Fruit fly).
C Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
C Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
C Ephydroidea; Drosophilidae; Drosophila.
X NCBI_TaxID=7227;
X [1]
N SEQUENCE FROM N.A.
C STRAIN=Berkley;
X MEDLINE=20196006; PubMed=10731132;
A Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
A Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
A George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
A Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
A Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,

```

```

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Beeson K.M., Basu A., Baxendale J., Bayraktarglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Fezra C., Ferrera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorteli J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guerin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH080.1-RESEARCH080.8(2002).
CC -1- SIMILARITY: Belongs to the UF0279 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE003800; AAF5714.1; -.
CC EMBL; AY084100; AAL89838.1; -.
CC FlyBase; FBgn0034327; CGI4505.
CC InterPro; IPR007967; DUF727.
CC Pfam; PF05503; DUF727; 1.
CC SEQUENCE 119 AA; 13534 MW; 83FA23FCCCE389AA CRC64;
Y Query Match 95.0%; Score 19; DB 1; Length 119;
Best Local Similarity 100.0%; Pred.No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 VAEF 5
|
|
|
33 VAEF 36
Db
RESULT 12
GTR_PIG STANDARD; PRT; 120 AA.
ID GTR2_PIG
AC O62786;
DT 28-FEB-2003 (Rel. 41, Created)

```

28-FEB-2003 (Rel. 41, Last sequence update)  
28-FEB-2003 (Rel. 41, Last annotation update)  
Solute carrier family 2, facilitated glucose transporter, member 2  
(Glucose transporter type 2, liver) (Fragment).  
SLC2A2 OR GLUT2.

Sus scrofa (Pig).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
NCBI\_TaxID=9823;  
[1]

SEQUENCE FROM N.A.  
Canty J.M., Young R.F., Fallavollita J.A.;  
Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
-!- FUNCTION: Facilitative glucose transporter. This isoform likely  
mediates the bidirectional transfer of glucose across the plasma  
membrane of hepatocytes and is responsible for uptake of glucose  
by the beta cells; may comprise part of the glucose-sensing  
mechanism of the beta cell. May also participate with the  
Na(+)/glucose cotransporter in the transcellular transport of  
glucose in the small intestine and kidney (By similarity).  
-!- SUBCELLULAR LOCATION: Integral membrane protein.  
-!- SIMILARITY: Belongs to the sugar transporter family. Glucose  
transporter subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF054835; AAC12737.1; -.  
InterPro; IPR007114; MFS.  
InterPro; IPR005828; Sub transporter.  
InterPro; IPR005829; Sug transporter.  
InterPro; IPR003663; Sug transprt.  
Pfam; PF00083; sugax tr; 1.  
PRINTS; PR00171; SUGTRNSPT.  
PROSITE; PS00950; MFS; 1.  
PROSITE; PS00216; SUGAR TRANSPORT 1; PARTIAL.  
PROSITE; PS00217; SUGAR TRANSPORT 2; PARTIAL.

Transmembrane; Sugar transporter; Transport; Multigene family.

NON TER 1 1  
TRANSMEM 2 22 9 (POTENTIAL).  
DOMAIN 23 26 EXTRACELLULAR (POTENTIAL).  
TRANSMEM 27 47 10 (POTENTIAL).  
DOMAIN 48 56 CYTOPLASMIC (POTENTIAL).  
TRANSMEM 57 77 11 (POTENTIAL).  
DOMAIN 78 84 EXTRACELLULAR (POTENTIAL).  
TRANSMEM 85 105 12 (POTENTIAL).  
DOMAIN 106 120 CYTOPLASMIC (POTENTIAL).  
NON TER 120 120  
SEQUENCE 120 AA; 13503 MW; DSF73168DBF03203 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 120;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||  
46 VAEF 49

35ULT 13  
19E\_PYRAE STANDARD; PRT; 122 AA.  
Q8ZY02;  
28-FEB-2003 (Rel. 41, Created)  
28-FEB-2003 (Rel. 41, Last sequence update)  
28-FEB-2003 (Rel. 41, Last annotation update)  
50S ribosomal protein L18e.  
RPL18E OR PAE0672.

Pyrobaculum aerophilum.  
Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
Thermoproteaceae; Pyrobaculum.  
NCBI\_TaxID=13773;  
[1]

SEQUENCE FROM N.A.  
STRAIN=IM2 / ATCC 51768 / DSM 7523;  
MEDLINE=21664397; PubMed=11792869;  
Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,  
Miller J.H.;  
RA "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum

aerophilum".  
RT aerophilum".  
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).

-!- SIMILARITY: Belongs to the L18E family of ribosomal proteins.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE009779; AAL62941.1; -.

HMAP; MF 00329; -; 1.  
InterPro; IPR001196; Ribosomal L15.  
InterPro; IPR000039; Ribosomal\_L18e.  
Pfam; PF00256; L15; 1.  
PROSITE; PS01106; RIBOSOMAL\_L18E; FALSE NEG.  
KW Ribosomal protein; Complete proteome.  
SEQUENCE 122 AA; 88F3DB732C4E0394 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 122;

Best Local Similarity 100.0%; Pred. No. 2.2e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0;

QY 2 VAEF 5  
|||  
35 VAEF 38

RESULT 14

Y670\_PASMU

ID Y670\_PASMU STANDARD; PRT; 124 AA.

AC Q9CWF0;  
DC 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical protein PM0670 precursor.  
GN PM0670.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Pasteurella.  
OX NCBI\_TaxID=747;  
[1]

SEQUENCE FROM N.A.

STRAIN=PM70;

RA MEDLINE=21145866; PubMed=11248100;

RA May B.J., Zhang Q., Li L.I., Paustian M.L., Whittam T.S., Kapur V.;

EL "Complete genomic sequence of Pasteurella multocida Pm70.";

Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

-!- SIMILARITY: Belongs to the cytochrome b562 family.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE006103; AAK02754.1; -.  
PIRSF; PIRSF0000029; Cytochrome\_b562; 1.

W Hypothetical protein; Signal; Complete proteome.  
T SIGNAL 1 23 POTENTIAL.  
T CHAIN 24 124 HYPOTHETICAL PROTEIN PM0670.  
Q SEQUENCE 124 AA; 13746 MW; D7B2B485C7B51B9A CRC64;

Query Match 95.0%; Score 19; DB 1; Length 124;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
|  
|  
|  
|  
b 45 VAEF 48

RESULT 15

LBP_MOUSE	STANDARD;	PRT;	127 AA.
D	PS1162;		
Y	01-OCT-1996 (Rel. 34, Created)		
Y	01-OCT-1996 (Rel. 34, Last sequence update)		
Y	10-OCT-2003 (Rel. 42, Last annotation update)		
Y	Gastrotropin (Gt) (Ileal lipid-binding protein) (ILBP).		
Y	FABP6 OR ILBP.		
Y	Mus musculus (Mouse).		
Y	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Y	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
Y	NCBI_TaxID=10090;		
Y	[1]		
Y	SEQUENCE FROM N.A.		
Y	STRAIN=DBA/2J, TISSUE=Liver;		
Y	MEDLINE=94375529; PubMed=8089185;		
Y	Crossman M.W., Haut S.M., Gordon J.I.;		
Y	"The mouse ileal lipid-binding protein gene: a model for studying		
Y	axial patterning during gut morphogenesis.";		
Y	J. Cell Biol. 126:1547-1564 (1994).		
Y	- FUNCTION: ILEAL PROTEIN WHICH STIMULATES GASTRIC ACID AND		
Y	PERSINOGEN SECRETION. SEEMS TO BE ABLE TO BIND TO BILE SALTS AND		
Y	BILIRUBINS.		
Y	- SUBCELLULAR LOCATION: Cytoplasmic.		
Y	- SIMILARITY: Belongs to the fatty-acid binding protein (FABP)		
Y	family.		
Y	-----		
Y	This SWISS-PROT entry is copyright. It is produced through a collaboration		
Y	between the Swiss Institute of Bioinformatics and the EMBL outstation		
Y	the European Bioinformatics Institute. There are no restrictions on its		
Y	use by non-profit institutions as long as its content is in no way		
Y	modified and this statement is not removed. Usage by and for commercial		
Y	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
Y	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
Y	-----		
Y	EMBL; U00938; AAC27352.1; -		
Y	PIR; A54797; A54797.		
Y	HSP; F10289; LFAL.		
Y	MGI; 96565; Fabbp6.		
Y	InterPro; IPR000463; Fatty_acid_BP.		
Y	InterPro; IPR000566; Lipogln_cytFABP.		
Y	Pfam; PF00061; lipocalin; 1.		
Y	PRINTS; PR00178; FATTYACIDBP.		
Y	PROSITE; PS00214; FABP; 1.		
Y	Transport; Lipid-binding; Acetylation.		
Y	INIT MET 0		
Y	MOD_RES 1		
Y	BY SIMILARITY		
Y	ACETYLATION (BY SIMILARITY).		
Y	SEQUENCE 127 AA; 14355 MW; 0690BDDDA09A322 CRC64;		
Y	-----		

```

RESULT 16
SPEH THEME STANDARD; PRG; 130 AA.
ID SPEH_THEME STANDARD; PRG; 130 AA.
AC Q9WZC3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50) (AdoMetDC)
DE (SAMDC) (Contains: S-adenosylmethionine decarboxylase beta chain; S-adenosylmethionine decarboxylase alpha chain).
DE SPEH OR TM0655.
GN Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
RC MDLLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickley E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Uutterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RA "Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -1- FUNCTION: Decarboxylation of S-adenosylmethionine provides the aminopropyl moiety required for spermidine biosynthesis from putrescine (By similarity)
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine = (S-deoxy-5-adenosyl)(3-aminopropyl) methylsulfonium salt + CO(2).
CC -1- COFACTOR: Pyruvoyl group (By similarity).
CC -1- SIMILARITY: Belongs to the prokaryotic AdoMetDC family. Subfamily 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001739; AAD35739.1; -.
CC PIR; D72348; D72348.
CC TIGR; TM0655; -.
CC HAMAP; MF_00464; -. 1.
CC InterPro; IPR003826; SAMDC.
CC Pfam; PF02675; AdoMetDC; 1.
CC Spermidine biosynthesis; Lyase; Decarboxylase; Zymogen; Pyruvate; Complete proteome.
CC CHAIN 1 62 S-ADENOSYLMETHIONINE DECARBOXYLASE BETA CHAIN (BY SIMILARITY).
CC CHAIN 63 130 S-ADENOSYLMETHIONINE DECARBOXYLASE ALPHA CHAIN (BY SIMILARITY).
CC SITE 62 63 CLEAVAGE (NONHYDROLYTIC) (BY SIMILARITY).
CC MOD_RES 63 63 CONVERTED TO A PYRUVOYL GROUP (BY SIMILARITY).
CC SEQUENCE 130 AA; 14785 MW; 7659FE20A2019928 CRC64;
Query Match 95.0%; Score 19; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 VAEF 5
DB 9 VAEF 12

```

RESULT 17  
Y194 PYRAB  
ID\_Y194 PYRAB STANDARD; PRT; 131 AA.

91 VAEF 94

Q9V280;  
16-OCT-2001 (Rel. 40, Last sequence update)  
16-OCT-2001 (Rel. 42, Last annotation update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
HYPOTHETICAL UPF0146 protein PYAB01940.  
PYAB01940 OR PAB2224.  
Pyrococcus abyssi.  
Archaea; Euryarchaeota; Thermococcales; Thermococcaceae;  
Pyrococcus.  
NCBI\_TaxID=29292;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=GES / Oreay;  
MEDLINE=22511545; PubMed=12622808;  
Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,  
Poch O., Priour D., Querellou J., Ripp R., Thierry J.-C.,  
van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.,  
"An integrated analysis of the genome of the hyperthermophilic  
archaeon Pyrococcus abyssi."  
Mol. Microbiol. 47:1495-1512(2003).  
-!- SIMILARITY: Belongs to the UPF0146 family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----  
EMBL; AJ248283; CAB49118.1; --  
PIR; G75208; G75208.  
HMAP; MF\_00341; --; 1.  
InterPro; IPR005353; UPF0146.  
Pfam; PF03686; UPF0146; 1.  
PIRSF; PIRSF016725; UCP016725; 1.  
ProDom; PD021130; UPF0146; 1.  
HYPOTHETICAL PROTEIN; Complete proteome.  
SEQUENCE 131 AA; 14629 MW; D13F378187P932C3 CRC64;  
-----  
Query Match 95.0%; Score 19; DB 1; Length 131;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 VAEF 5  
4 VAEF 7  
-----  
RESULT 18  
RISB\_HALM1 STANDARD; PRT; 133 AA.  
O9HRM5;  
16-OCT-2001 (Rel. 40, Created)  
16-OCT-2001 (Rel. 40, Last sequence update)  
28-FEB-2003 (Rel. 41, Last annotation update)  
6-7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.9) (DMRL synthase)  
(lumazine synthase) (Riboflavin synthase beta chain).  
RIBH OR RIBE OR VNG0630G.  
X Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).  
Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
Halobacteriaceae; Halobacterium.  
NCBI\_TaxID=64091;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=20504483; PubMed=11016950;  
Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,  
Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
Leitthausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
Tsenbarger T.A., Peck R.F., Pohlshocher M., Spudich J.L., Jung K.-H.,  
Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,

EBhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;  
"Genome sequence of Halobacterium species NRC-1."  
Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
-!- FUNCTION. Riboflavin synthase is a bifunctional enzyme complex  
catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-  
ribityl-amino-2,4-(1H,3H)-pyrimidinedione and L-3,4-dihydroxy-2-  
butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit  
catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-amino-  
2,4-(1H,3H)-pyrimidinedione with L-3,4-dihydroxy-2-butanone-4-  
phosphate yielding 6,7-dimethyl-8-lumazine (By similarity).  
-!- CATALYTIC ACTIVITY: 2 6,7-dimethyl-8-(1-D-ribityl)lumazine =  
riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.  
-!- PATHWAY: Riboflavin biosynthesis; last step.  
-!- SIMILARITY: Belongs to the DMRL synthase family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----  
EMBL; AE005011; AAC19133.1; --  
PIR; AB4221; AB4221.  
HMAP; MF\_00178; --; 1.  
InterPro; IPR002180; DMRL synthase.  
Pfam; PF00885; DMRL synthase; 1.  
ProDom; PD003664; DMRL synthase; 1.  
RIBOFLAVIN BIOSYNTHESIS; Transferase; Complete proteome.  
SEQUENCE 133 AA; 13719 MW; B9734DC24660D6CD CRC64;  
-----  
Query Match 95.0%; Score 19; DB 1; Length 133;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 VAEF 5  
8 VAEF 11  
-----  
RESULT 19  
YBGC\_ECOLI STANDARD; PRT; 134 AA.  
ID\_YBGC\_ECOLI  
AC P08939;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Protein ybgC.  
GN YBGC OR B0736 OR C0815 OR Z0904 OR ECS0771 OR SF0561 OR S0574.  
OS Escherichia coli.  
OS Escherichia coli O5.  
OS Escherichia coli O157:H7, and  
OS Shigella flexneri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 217992, 83334, 523;  
[1]  
SEQUENCE FROM N.A.  
RP SPECIES=E.coli;  
RC MEDLINE=8722192; PubMed=3294803;  
RA Sun T.-P., Webster R.E.;  
RT "Nucleotide sequence of a gene cluster involved in entry of E colicins  
and single-stranded DNA of infecting filamentous bacteriophages into  
Escherichia coli."  
RL J. Bacteriol. 159:2667-2674(1987).  
[2]  
SEQUENCE FROM N.A.  
RP SPECIES=E.coli; STRAIN=K12 / MG1655;  
RC MEDLINE=97426617; PubMed=9278503;  
RX Blattner F.R., Plunkett G. III, Bloch C.A., Ferna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
A Mau B., Shao Y.;  
"The complete genome sequence of *Escherichia coli* K-12.";  
Science 277:1453-1474 (1997).  
[3]  
SEQUENCE FROM N.A.  
SPECIES=E.coli; STRAIN=K12;  
MEDLINE=97061202; PubMed=8905232;  
X Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
A Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
A Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
A Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
A Yano M., Horiuchi T.,  
A "A 718-Kb DNA sequence of the *Escherichia coli* K-12 genome  
corresponding to the 12.7-28.0 min region on the linkage map.";  
DNA Res. 3:137-155 (1996).  
[4]  
SEQUENCE FROM N.A.  
SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;  
C MEDLINE=22388234; PubMed=12471157;  
X Wei R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
A Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
A Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
A Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
T "Extensive mosaic structure revealed by the complete genome sequence  
of uropathogenic *Escherichia coli* K12";  
Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).  
[5]  
SEQUENCE FROM N.A.  
SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;  
C MEDLINE=21074935; PubMed=11206551;  
X Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
A Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
A Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
A Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
A Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
A Welch R.A., Blattner F.R.;  
T "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7";  
Nature 409:529-533 (2001).  
[6]  
SEQUENCE FROM N.A.  
SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;  
C MEDLINE=21156231; PubMed=11258796;  
X Hayaishi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,  
A Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
A Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
A Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
T "Complete genome sequence of enterohemorrhagic *Escherichia coli*  
O157:H7 and genomic comparison with a laboratory strain K-12.";  
DNA Res. 8:11-22 (2001).  
[7]  
SEQUENCE OF 1-20 FROM N.A.  
SPECIES=E.coli; STRAIN=K12 / MGL1655;  
C Kim K., Allen E., Araujo R., Aparicio A.M., Botstein D.,  
A Cherry M., Chung E., Dietrich F., Duncan M., Federspiel N.,  
A Kaiman S., Komp C., Lashkari D., Lew H., Lin D., Namath A.,  
A Osiner P., Davis R.;  
L Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
[8]  
SEQUENCE FROM N.A.  
SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;  
C MEDLINE=22272406; PubMed=12384590;  
X Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
A Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
A Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Hou S.,  
A Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
A Yu J.;  
T "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity  
through comparison with genomes of *Escherichia coli* K12 and O157";  
Nucleic Acids Res. 30:4432-4441 (2002).  
[9]  
SEQUENCE FROM N.A.  
SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;  
C MEDLINE=22590274; PubMed=12704152; Venkatesan M.M., Deng W.,  
X Wei J., Goldberg M.B., Burland V., Plunkett G. III, Rose D.J., Darling A.,  
A Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,  
A Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,  
A Schwartz D.C., Blattner F.R.;  
T "Complete genome sequence and comparative genomics of *Shigella*  
flexneri serotype 2a strain 2457T";  
Infect. Immun. 71:2775-2786 (2003).  
[10]  
IDENTIFICATION BY MASS SPECTROMETRY.  
SPECIES=E.coli;  
C MEDLINE=99420866; PubMed=10493123;  
X Fountoulakis M., Takacs M.-F., Berndt P., Langen H., Takacs B.;  
A "Enrichment of low abundance proteins of *Escherichia coli* by  
hydroxyapatite chromatography";  
Electrophoresis 20:2181-2195 (1999).  
C -1- SIMILARITY: BELONGS TO THE 4-HYDROXYBENZOVYL-COA THIOESTERASE  
FAMILY. STRONG, TO H.INFLUENZAE HI0386.  
C This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
C  
C EMBL: M16489; AAA83918.1; -  
C EMBL: AE000177; AAC73810.1; -  
C EMBL: D90713; BAA35402.1; -  
C EMBL: AE016757; AAN79288.1; -  
C EMBL: AE005252; AAG55072.1; -  
C EMBL: AP002553; BAB34194.1; -  
C EMBL: U30934; AAA74398.1; -  
C EMBL: AE015086; AAN42205.1; -  
C EMBL: AE015979; AAP16078.1; -  
C PIR: A35980; WMEC15.  
C PIR: C90725; C90725.  
C PIR: D85576; D85576.  
C Ecogene; EGI1110; ybgC.  
C InterPro; IPR008272; 4HBCOA\_thioest\_AS.  
C InterPro; IPR006684; 4HBCOA\_thioestise.  
C InterPro; IPR006683; Thioestr\_supf.  
C Pfam; PF03061; 4HET; 1  
C TIGRFAMs; TIGR00051; TIGR00051; 1  
C PROSITE; PS01328; 4HBCOA\_THIOESTERASE; 1.  
C Hydrolase; Complete proteome.  
C ACT SITE 18 18 BY SIMILARITY.  
C ACT SITE 18 18  
C SEQUENCE 134 AA; 15562 MW; C44582B6EC3BE989 CRC64;  
Query Match 95.0%; Score 19; DB 1; Length 134;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 129 VAEF 132  
RESULT 20  
RISB METH STANDARD; PRT; 139 AA.  
ID RISB METH  
AC 027473;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 6,7-dimethyl-8-ribitylumazine synthase (EC 2.5.1.9) (DMRL synthase)  
DE (lumazine synthase) (Riboflavin synthase beta chain).  
GN RIBH OR MTK1390.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
OC Methanobacteriaceae; Methanothermobacter.

NCBI\_TaxID=187420;

[1]

SEQUENCE FROM N.A.

STRAIN=Delta H;

MEDLINE=98037514; PubMed=9371463;

Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
Alredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
Harrison D., Hough L., Kagle P., Lum W., Pothier B., Qiu D.,  
Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
Daniels C.J., Mao J.-I., Rice P., Noellings J., Reeve J.N.;  
"Complete genome sequence of *Methanobacterium thermoautotrophicum*  
deltaH: functional analysis and comparative genomics.";  
J. Bacteriol. 179:1735-1755(1997).

-|- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex  
catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-  
ribityl-amino-2,4(1H,3H)-pyrimidinone and L-3,4-dihydroxy-2-  
butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit  
catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-amino-  
2,4(1H,3H)-pyrimidinone with L-3,4-dihydroxy-2-butanone-4-  
phosphate yielding 6,7-dimethyl-8-lumazine (by similarity).

-|- CATALYTIC ACTIVITY: 2 6,7-dimethyl-8-(1-D-ribityl)lumazine =  
riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.

-|- PATHWAY: Riboflavin biosynthesis; last step.

-|- SIMILARITY: Belongs to the DMRL synthase family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE000902; AAB85867.1; -

PIR; A69052; A69052.

HSNP; P11998; IRVV.

HAWAP; MF 00178; -; 1.

InterPro; IPR002180; DMRL synthase.

Pfam; PF00885; DMRL synthase; 1.

Pfam; PF003664; DMRL synthase; 1.

ProDom; PD003664; DMRL synthase; 1.

TIGRFAMs; TIGR00114; ribH; 1.

Riboflavin biosynthesis; Transferrase; Complete proteome.

SEQUENCE 139 AA; 15464 MW; B640D8D12EE2A862 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 139;

Best Local Similarity 100.0%; Pred. No. 2.5e+02; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5

||||

10 VAEF 13

SUIT 21

86 ARCFU

Y586 ARCFU STANDARD; PRT; 140 AA.

029669;

16-OCT-2001 (Rel. 40, Created)

16-OCT-2001 (Rel. 40, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)

Hypothetical protein AF0586.

AF0586.

Archaeoglobus fulgidus.

Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;

Archaeoglobaceae; Archaeoglobus.

NCBI\_TaxID=2234;

[1]

SEQUENCE FROM N.A.

STRAIN=VC-16 / DSM 4304 / ATCC 49558;

MEDLINE=98049343; PubMed=9389475;

Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,  
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
RA Overbeek R., Gocayne J.D., Weidman J.P., McDonald L., Utterback T.,  
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,  
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
RA Venter J.C.;

RT "The complete genome sequence of the hyperthermophilic, sulphate-

reducing archaeon *Archaeoglobus fulgidus*.";

RL Nature 390:364-370(1997).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; AE001064; AAB90658.1; -

DR PIR; B69323; B69323.

DR TIGR; AF0586; -

KW Hypothetical protein; Complete proteome.

SEQUENCE 140 AA; 16175 MW; D3B1PE3A62A05901 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 140;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5

||||

65 VAEF 68

Db

RESULT 22

RISB ARCFU

ID RISB ARCFU STANDARD; PRT; 143 AA.

AC 028152;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.9) (DMRL synthase)

DE (lumazine synthase) (Riboflavin synthase beta chain).

GN RIBH OR RIBE OR AF2128.

OS Archaeoglobus fulgidus.

OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;

OC Archaeoglobaceae; Archaeoglobus.

OX NCBI\_TaxID=2234;

RN [1]

SEQUENCE FROM N.A.

STRAIN=VC-16 / DSM 4304 / ATCC 49558;

MEDLINE=98049343; PubMed=9389475;

Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,

RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,

RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,

RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,

RA Overbeek R., Gocayne J.D., Weidman J.P., McDonald L., Utterback T.,

RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,

RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,

RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,

RA Venter J.C.;

RT "The complete genome sequence of the hyperthermophilic, sulphate-

reducing archaeon *Archaeoglobus fulgidus*.";

RL Nature 390:364-370(1997).

CC -|- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex

CC catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-

CC ribityl-amino-2,4(1H,3H)-pyrimidinone and L-3,4-dihydroxy-2-

CC butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit





```

CC -!- SIMILARITY: Contains 3 EF-hand calcium-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB036744; BAA95412.1; -
CC DR EMBL; AK009956; BAB26608.1; -
CC DR EMBL; AK009664; BAB26425.1; -
CC DR EMBL; AK003648; BAB22914.1; -
CC DR HSSP; P02593; 1CDM.
CC DR MGD; MGI:1931464; Calm4
CC DR InterPro; IPR002048; EF-hand.
CC DR InterPro; IPR001125; Recoverin.
CC DR Pfam; PF00036; ehand; 4.
CC DR PRINTS; PRO0450; RECOVERIN.
CC DR ProDom; PD000012; EF-hand; 2.
CC DR SMART; SM00054; Efh; 4.
CC DR PROSITE; PS00018; EF-hand; 3.
CC KW Calcium-binding; Repeat.
CC FT CA_BIND 21 32 EF-HAND 1 (POTENTIAL).
CC FT CA_BIND 57 68 EF-HAND 2 (POTENTIAL).
CC FT CA_BIND 93 104 EF-HAND 3 (POTENTIAL).
CC FT CONFLICT 9 9 E -> V (IN REF. 2; BAB22914).
CC FT CONFLICT 124 124 M -> V (IN REF. 2; BAB22914).
CC FT CONFLICT 146 148 VEN -> I (IN REF. 1).
CC SQ SEQUENCE 148 AA; 16767 MW; 2AE22BC738AD2F9D CRC64;
CC
CC Query Match 95.0%; Score 19; DB 1; Length 148;
CC Best Local Similarity 100.0%; Pred. No. 2.6e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 2 VAEF 5
CC DB 10 VAEF 13
CC
CC RESULT 25
CC YORX_PYRO STANDARD; PRT; 148 AA.
CC ID YORX_PYRO
CC AC P0259;
CC DT 01-FEB-1991 (Rel. 17, Created)
CC JT 01-FEB-1991 (Rel. 17, Last sequence update)
CC JT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Hypothetical protein in GAPDH 3'region (ORF X) (Fragment).
CC DS Pyrococcus woesei.
CC SC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CC NCBI_TaxID=2262;
CC RZ SEQUENCE FROM N.A.
CC RP STRAIN=DSM 3773;
CC RX MEDLINE=90330536; PubMed=2165475;
CC ZA Zwickl P., Fabry S., Bogaedain C., Haas A., Hensel R.;
CC XT "Glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic
CC XT archaeobacterium Pyrococcus woesei: characterization of the enzyme,
CC XT cloning and sequencing of the gene, and expression in Escherichia
CC XT coli.";
CC RL J. Bacteriol. 172:4329-4338(1990).
CC DR PIR; S10655; S10655.
CC KW Hypothetical protein.
CC FT NON_TER 1
CC SQ SEQUENCE 148 AA; 15256 MW; 8C7E51A072DF5D50 CRC64;
CC
CC Query Match 95.0%; Score 19; DB 1; Length 148;
CC Best Local Similarity 100.0%; Pred. No. 2.6e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC YZ 2 VAEF 5

```

```

DB 115 VAEF 118
DB
RESULT 26
RISB_PYRAE STANDARD; PRT; 150 AA.
ID RISB_PYRAE
AC Q82TE3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.9) (DMRL synthase)
DE (Lumazine synthase) (Riboflavin synthase beta chain).
DE RIBH OR PAE3296.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RX Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RA "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
CC -!- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex
CC catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-
CC ribityl-amino-2,4(1H,3H)-pyrimidinone and L-3,4-dihydroxy-2-
CC butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit
CC catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-
CC amino-2,4(1H,3H)-pyrimidinone with L-3,4-dihydroxy-2-
CC butanone-4-phosphate yielding 6,7-dimethyl-8-lumazine (By
CC similarity)
CC -!- CATALYTIC ACTIVITY: 2 6,7-dimethyl-8-(1-D-ribityl)lumazine =
CC riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.
CC -!- PATHWAY: Riboflavin biosynthesis; last step.
CC -!- SIMILARITY: Belongs to the DMRL synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE009923; AAL64819.1; -
CC DR HAVAP; MF 00178; -; 1
CC DR InterPro; IPR002180; DMRL_synthase.
CC DR Pfam; PF00885; DMRL_synthase; 1.
CC DR ProDom; PD003664; DMRL_synthase; 1.
CC DR TIGRFAMs; TIGR00114; RibH; 1.
CC KW Riboflavin biosynthesis; Transferase; Complete proteome.
CC SQ SEQUENCE 150 AA; 16474 MW; 73751C1363A8AD8D CRC64;
CC
CC Query Match 95.0%; Score 19; DB 1; Length 150;
CC Best Local Similarity 100.0%; Pred. No. 2.6e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 2 VAEF 5
CC DB 8 VAEF 11
CC
CC RESULT 27
CC SODC_FRIGL STANDARD; PRT; 152 AA.
ID SODC_FRIGL
AC P11418;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

```



DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).

OS Prionace glauca (Blue shark).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;

OC Elasmobranchii; Galeomorphi; Galeoidea; Carcharhiniformes;

OC Carcharhinidae; Prionace.

OX NCBI\_TaxID=7815;

RN [1]

RP SEQUENCE.

RX MEDLINE=89290032; PubMed=2500367;

RA Calabrese L., Politicelli F., O'Neill P., Gallieri A., Barra D.,

RA Schinina M.E., Bossa F.;

RT "Substitution of arginine for lysine 134 alters electrostatic

RT parameters of the active site in shark Cu,Zn superoxide dismutase.";

RL FEBS Lett. 250:49-52(1989).

CC -!- FUNCTION: Destroys radicals which are normally produced within the

CC cells and which are toxic to biological systems.

CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).

CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit.

CC -!- SUBUNIT: Homodimer.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.

DR PIR; S04623; S04623.

DR HSSP; P00442; 1CEJ.

DR InterPro; IPR001424; SOD\_CU\_ZN.

DR Pfam; PF00080; sdcu; 1.

DR PRINTS; PR00068; CUZNDISMTASE.

DR ProDom; PD000469; SOD\_CU\_ZN\_1; 1.

DR PROSITE; PS00087; SOD\_CU\_ZN\_1; 1.

DR PROSITE; PS00332; SOD\_CU\_ZN\_2; 1.

KW Antioxidant; Oxidoreductase; Metal-binding; Copper; Zinc.

FT METAL 44 44

FT METAL 46 46 COPPER.

FT METAL 61 61 COPPER AND ZINC.

FT METAL 69 69 ZINC.

FT METAL 78 78 ZINC.

FT METAL 81 81 ZINC.

FT METAL 118 118 COPPER.

FT DISULFID 55 144 BY SIMILARITY.

SQ SEQUENCE 152 AA; 15840 MW; 6617642A4F23CSAE CRC64;

Query Match 95.0%; Score 19; DB 1; Length 152;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5

Db 92 VAEF 95

RESULT 28

ID RISE\_SULSO STANDARD; PRT; 154 AA.

AC Q990B5;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE 6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.9) (DMRL synthase)

DE (lumazine synthase) (Riboflavin synthase beta chain).

GN RIBH OR SSO0400.

OS Sulfolobus solfataricus.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI\_TaxID=2287;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=ATCC 35092 / DSM 1617 / B2;

RX MEDLINE=21332296; PubMed=11427726;

RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,

RA Aways M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,

RA De Moors A., Traus G., Fletcher C., Gordon P.M.K.,

RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,

RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,

RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,

RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;

RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).

CC -!- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex

CC catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-

CC ribityl-amino-2,4(1H,3H)-pyrimidinedione and L-3,4-dihydroxy-2-

CC butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit

CC catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-

CC amino-2,4(1H,3H)-pyrimidinedione with L-3,4-dihydroxy-2-

CC butanone-4-phosphate yielding 6,7-dimethyl-8-lumazine (By

CC similarity).

CC -!- CATALYTIC ACTIVITY: 2 6,7-dimethyl-8-(1'-D-ribityl)lumazine =

CC riboflavin + 4-(1'-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.

CC -!- PATHWAY: Riboflavin biosynthesis; last step.

CC -!- SIMILARITY: Belongs to the DMRL synthase family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AE006673; AAK40729.1; -.

DR PIR; B90184; B90184.

DR HAMAP; MF 00178; -; 1.

DR InterPro; IPR002180; DMRL synthase.

DR Pfam; PF00885; DMRL synthase; 1.

DR ProDom; PD003664; DMRL synthase; 1.

DR TIGRfams; TIGR00114; rLBH; 1.

KW Riboflavin biosynthesis; Transferase; Complete proteome.

SQ SEQUENCE 154 AA; 17247 MW; ADPA9B6A2C723210 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 154;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5

Db 12 VAEF 15

RESULT 29

ID PCP\_HABIN STANDARD; PRT; 155 AA.

AC P10325;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Outer membrane lipoprotein PCP precursor (15 kDa lipoprotein) (PAL

DE cross-reacting lipoprotein).

GN PCP OR LPP OR H11579.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Haemophilus.

OX NCBI\_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88115138; PubMed=2828309;

RA Daich R.A., Metcalf B.J., Finn C.W., Farley J.B., Green B.A.;

RT "Cloning of genes encoding a 15,000-dalton peptidoglycan-associated

RT outer membrane lipoprotein and an antigenically related 15,000-dalton

RT protein from Haemophilus influenzae.";

RL J. Bacteriol. 170:489-498(1988).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=RD / RW20 / ATCC 51907;

RX MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

RA Karlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 Fine L.D., Pritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 Venter J.C.;  
 "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 Rd.";  
 Science 269:496-512(1995).  
 -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid  
 anchor.  
 -1- SIMILARITY: TO E.COLI AND S.TYPHIMURIUM SLVB AND TO  
 Y.ENTEROCOLITICA PCP.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M18877; AAA24938.1; -;  
 EMBL; U32832; AAC23228.1; -;  
 PIR; I64130; I64130.

TIGR; H11579; -;  
 InterPro; IPR000437; Prok lipoprot. S.  
 PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.  
 Outer membrane; Lipoprotein; Signal; Complete proteome; Palmitate.  
 SIGNAL 1 18  
 CHAIN 19 155 OUTER MEMBRANE LIPOPROTEIN PCP.  
 LIPID 19 19 N-palmitoyl cysteine.  
 LIPID 19 19 S-diacylglycerol cysteine.  
 CONFLICT 135 143 CSUVAERFV -> VAGKRVRI (IN REF. 1).  
 SEQUENCE 155 AA; 15425 MW; D7880327FCFC0C985 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 155;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
 ||||  
 138 VAEF 141

SULT 30  
 P\_TAROF STANDARD; PRT; 157 AA.  
 O49065;

30-MAY-2000 (Rel. 39, Created)  
 30-MAY-2000 (Rel. 39, Last sequence update)  
 15-MAR-2004 (Rel. 43, Last annotation update)  
 Root allergen protein (RAP).  
 Taraxacum officinale (Common dandelion).  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;  
 Taraxacum.  
 NCBI\_TaxID=50225;  
 [1]

SEQUENCE FROM N.A.

TISSUE=Root;  
 Xu X.-Y., Bewley J.D., Greenwood J.S.;  
 Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 -1- ALLERGEN: Causes an allergic reaction in human.  
 -1- SIMILARITY: Belongs to the BetVI family.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF036931; AAB92255.1; -;  
 DR HSP; O24248; 1E09.  
 DR InterPro; IPR000916; Bet v.I.  
 DR Pfam; PF00407; Bet v.I; 1  
 DR PRINTS; PR00634; BETALLERGEN.  
 DR PROSITE; PS00451; PATHOGENESIS BETVI; 1.  
 KW Allergen; Plant defense; Pathogenesis-related protein.  
 SQ SEQUENCE 157 AA; 17040 MW; 5892AB8593A8A7E0 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 ||||  
 Db 3 VAEF 6

RESULT 31

RISB SULTO STANDARD; PRT; 157 AA.  
 ID RISB SULTO  
 AC Q975N5;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE 6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.9) (DMRL synthase)  
 DE (lumazine synthase) (Riboflavin synthase beta chain).  
 GN RIBH CR ST0394.  
 OS Sulfolobus tokodaii.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 Sulfolobus.  
 OC NCBI\_TaxID=111955;  
 OX  
 RN [1]  
 SEQUENCE FROM N.A.  
 RP STRAIN=JCM 10545 / 7;  
 RC MEDLINE=21456156; PubMed=11572479;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
 Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,  
 Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
 Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
 Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
 Oshima T., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic thermophilic  
 Crenarchaeon, Sulfolobus tokodaii strain7.";  
 RL DNA Res. 8:123-140(2001).

-1- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex  
 catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-  
 ribityl-amino-2,4-(1H,3H)-pyrimidinedione and L-3,4-dihydroxy-2-  
 butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit  
 catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-  
 amino-2,4-(1H,3H)-pyrimidinedione with L-3,4-dihydroxy-2-  
 butanone-4-phosphate yielding 6,7-dimethyl-8-lumazine (By  
 similarity).

-1- CATALYTIC ACTIVITY: 2 6,7-dimethyl-8-(1-D-ribityl)lumazine =  
 riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.  
 -1- PATHWAY: Riboflavin biosynthesis; last step.

-1- SIMILARITY: Belongs to the DMRL synthase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AP000982; BAB65375.1; -;

HANAP; MF\_00178; -;

InterPro; IPR002180; DMRL synthase.

Pfam; PF00885; DMRL synthase; 1.

ProDom; PD003664; DMRL synthase; 1.

TIGRFAMs; TIGR00114; ribH; 1.

W Riboflavin biosynthesis; Transferase; Complete proteome.  
 2 SEQUENCE 157 AA; 17501 MW; 529558077E164A18 CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Y 2 VAEF 5  
 b 14 VAEF 17  
 RESULT 32  
 D SG55 DROME STANDARD; PRT; 163 AA.  
 C P07701; Q9VE1S;  
 T 01-APR-1988 (Rel. 07, Created)  
 T 01-APR-1988 (Rel. 07, Last sequence update)  
 T 10-OCT-2003 (Rel. 42, Last annotation update)  
 E Salivary glue protein sgs-5 precursor.  
 N SG55 OR CG7596.  
 S Drosophila melanogaster (Fruit fly).  
 C Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 C Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 C Ephydroidea; Drosophilidae; Drosophila.  
 X NCBI\_TaxID=7227;  
 N [1]  
 P SEQUENCE FROM N.A.  
 C STRAIN=Oregon-R;  
 C MEDLINE=87086754; PubMed=3098981;  
 A Shore E.M., Guild G.M.;  
 T "Larval salivary gland secretion proteins in Drosophila structural  
 analysis of the sgs-5 gene."  
 J. Mol. Biol. 190:149-158(1985).  
 N [2]  
 P SEQUENCE FROM N.A.  
 C STRAIN=Beckley;  
 C MEDLINE=20196006; PubMed=10731132;  
 A Adams M.D., Gelniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 A Amanatides P.G., Scher S.E., Li P.W., Hoekins R.A., Galle R.F.,  
 A George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 A Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 A Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 A Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 A Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 A Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 A Beeson K.Y., Beeson P.V., Berland B.P., Bhandari D., Bolshakov S.,  
 A Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 A Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 A Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 A de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 A Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 A Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 A Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 A Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 A Harris N.B., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 A Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 A Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 A Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 A Lasok P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 A Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson D.,  
 A Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 A Mount S.M., Moy M., Murphy B., Murphy L., Murthy D.M., Nelson D.B.,  
 A Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 A Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 A Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 A Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 A Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 A Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 A Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 A Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 A Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 A Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X04269; CA227820.1; -;  
 DR EMBL; AF003718; AF555436.1; -;  
 DR PIR; A24504; A24504.  
 DR FlyBase; FBgn0003375; Sgs5.  
 KW Signal.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 163 SALIVARY GLUE PROTEIN SGS-5.  
 SQ SEQUENCE 163 AA; 18821 MW; 3A476F1D3B06D864 CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 DB 114 VAEF 117  
 RESULT 33  
 ID UTRS YEAST STANDARD; PRT; 166 AA.  
 AC P32630;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE UTRS protein (unknown transcript 5 protein).  
 GN UTRS OR YEL035C OR SYGP-ORF27.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B-6441;  
 RX MEDLINE=94016558; PubMed=8411151;  
 RA Melnick L., Sherman F.;  
 RT "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,  
 RL of Saccharomyces cerevisiae share a common ancestry."  
 J. Mol. Biol. 233:372-388(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=97313264; PubMed=9169868;  
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,  
 RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,  
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,  
 RA Hunkeler-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,  
 RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,  
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V."  
 RL Nature 387:78-81(1997).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; L22173; AAA34936.1; -  
EMBL; S65964; AAD13970.1; -  
EMBL; S66120; AAB28442.1; -  
EMBL; U18779; AAB65007.1; -  
PIR; S50509; S50509.  
GermOnline; 139039; -  
SDD; S0000761; UTR5.  
CONFLICT 1 20 MSRYGKLVHYIIVHDDOR -> MRDSNVKISVFPFCALYN  
RGNTIN (IN REF. 1).  
SEQUENCE 166 AA; 19335 MW; A58E89DFA6D3DEA CRC64;  
Query Match 95.0%; Score 19; DB 1; Length 166;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 VAEF 5  
92 VAEF 95  
SULT 34  
B\_RHOBA  
SSB RHOBA STANDARD; PRT; 169 AA.  
P59932; Q7UKV3;  
15-MAR-2004 (Rel. 43, Created)  
15-MAR-2004 (Rel. 43, Last sequence update)  
15-MAR-2004 (Rel. 43, Last annotation update)  
Single-strand binding protein (SSB) (Helix-destabilizing protein).  
SSB OR R89917.  
Rhodospirillum rubrum.  
Bacteria; Planctomycetes; Planctomycetaceae; Planctomycetales;  
Planctomycetaceae; Pirellula.  
NCBI\_TaxID=117;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=1;  
MEDLINE=2735913; PubMed=12835416;  
Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,  
Schlesner H., Amann R., Reinhardt R.;  
"Complete genome sequence of the marine planctomycete Pirellula sp.  
strain 1".  
Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
-!- FUNCTION: This protein is essential for replication of the  
chromosome. It is also involved in DNA recombination and repair  
(By similarity).  
-!- SIMILARITY: Contains 1 SSB domain.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; BX294150; CAD76529.1; -  
PROSITE; PS50935; SSB; 1.  
DNA-binding; DNA repair; DNA replication; Complete proteome.  
DOMAIN 4 107  
DOMAIN 113 121 POLY-GLY.  
SEQUENCE 169 AA; 18108 MW; AA8F93E0FA51F287 CRC64;  
Query Match 95.0%; Score 19; DB 1; Length 169;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 VAEF 5  
28 VAEF 31  
RESULT 35

BFL1 MOUSE  
ID BFL1 MOUSE STANDARD; PRT; 172 AA.  
AC Q07450;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Bcl-2-related protein A1 (BFL-1 protein) (Hemopoietic-specific early  
response protein) (A1-A).  
GN BCL2A1 OR BCL2A1A OR BFL1 OR A1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CBA/J; TISSUE=Bone marrow;  
RX MEDLINE=93346743; PubMed=8345191;  
RA Lin E.Y., Orloffsky A., Berger M.S., Pryatowsky M.B.;  
"Characterization of A1, a novel hemopoietic-specific early-response  
gene with sequence similarity to bcl-2".  
J. Immunol. 151:1979-1988(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/Sv; TISSUE=Liver;  
RX MEDLINE=98307518; PubMed=9645611;  
RA Hatakeyama S., Hamasaki A., Negishi I., Loh D.Y., Sando F.,  
Nakayama K., Nakayama K.-I.;  
"Multiple gene duplication and expression of mouse bcl-2-related  
genes, A1".  
Int. Immunol. 10:631-637(1998).  
-!- FUNCTION: Retards apoptosis induced by IL-3 deprivation. May  
function in the response of hemopoietic cells to external signals  
and in maintaining endothelial survival during infection.  
-!- SUBCELLULAR LOCATION: Intracellular.  
-!- TISSUE SPECIFICITY: Expressed in hemopoietic tissues, including  
bone marrow, spleen and thymus.  
-!- INDUCTION: By granulocyte-macrophage colony-stimulating factor  
and LPS in macrophages.  
-!- SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.  
-!- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.  
-!- SIMILARITY: Belongs to the Bcl-2 family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; L16462; AAA16886.1; -  
EMBL; U23774; AAB97953.1; -  
EMBL; U23773; AAB97953.1; JOINED.  
PIR; I49449; I49449.  
HSPF; Q07817; IMAZ.  
MGD; MGI:102687; Bcl2a1a.  
InterPro; IPR000712; Bcl2 BH.  
InterPro; IPR002475; BCL2\_family.  
PFam; PF00452; Bcl-2; 1.  
SMART; SMC0337; BCL; 1.  
PROSITE; PS50062; BCL2\_FAMILY; 1.  
PROSITE; PS01080; BH1; 1.  
PROSITE; PS01258; BH2; 1.  
Apoptosis.  
KW DOVAIN 24 33 ALA/PRO-RICH.  
FT DOVAIN 77 97 BH1.  
FT DOVAIN 132 147 BH2.  
SQ SEQUENCE 172 AA; 19914 MW; 37AD35818E756488 CRC64;  
Query Match 95.0%; Score 19; DB 1; Length 172;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



the first processed in amphibia.":  
 J. Biol. Chem. 262:7901-7907(1987).  
 [2]  
 X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 MEDLINE=99167924; PubMed=10439069;  
 Ha Y., Shi D., Small G.W., Threl E.C., Allwell N.M.;  
 "Crystal structure of bullfrog M ferritin at 2.8 A resolution:  
 analysis of subunit interactions and the binuclear metal center.";  
 J. Biol. Inorg. Chem. 4:243-256(1999).  
 -!- FUNCTION: Ferritin is an intracellular molecule that stores iron  
 in a soluble, nontoxic, readily available form. The functional  
 molecule, which is composed of 24 chains, is roughly spherical and  
 contains a central cavity into which the polymeric ferric iron  
 core is deposited.  
 -!- MISCELLANEOUS: THERE ARE THREE TYPES OF FERRITIN SUBUNITS: L, M  
 AND H CHAINS IN AMPHIBIA.  
 -!- SIMILARITY: Belongs to the ferritin family.  
 -!- SIMILARITY: Contains 1 ferritin-like diiron domain.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; J02724; AAA49525.1; -;  
 PIR; C27805; C27805.  
 PDB; 1MFR; 22-JUN-99.  
 InterPro; IPR001519; Ferritin.  
 InterPro; IPR008331; Ferritin Dps.  
 InterPro; IPR008040; Ferritin\_like.  
 Pfam; PF00210; ferritin; 1.  
 ProDom; PD000971; Ferritin; 1.  
 PROSITE; PS00540; FERRITIN 1; 1.  
 PROSITE; PS00204; FERRITIN 2; 1.  
 PROSITE; PS09095; FERRITIN LIKE; 1.  
 Iron storage; Iron; Metal-binding; 3D-structure.  
 DOMAIN 7 156 FERRITIN-LIKE DIIRON.  
 METAL 24 24 IRON (BY SIMILARITY).  
 METAL 58 58 IRON (BY SIMILARITY).  
 METAL 59 59 IRON (BY SIMILARITY).  
 METAL 62 62 IRON (BY SIMILARITY).  
 METAL 104 104 IRON (BY SIMILARITY).  
 METAL 138 138 IRON (BY SIMILARITY).  
 METAL 138 138 IRON (BY SIMILARITY).  
 HELIX 11 39  
 TURN 41 43  
 TURN 46 73  
 HELIX 74 74  
 TURN 82 82  
 STRAND 93 120  
 HELIX 121 122  
 TURN 124 133  
 HELIX 134 134  
 TURN 135 154  
 TURN 155 159  
 TURN 161 170  
 TURN 171 171  
 QY SEQUENCE 176 AA; 20592 MW; A9F0F5BEB8584D46 CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Y 2 VAEF 5  
 48 VAEF 51  
 -----  
 RESULT 38  
 RL10\_THETN  
 ID \_RL10\_THETN STANDARD; PRT; 177 AA.

Q8R7U4;  
 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L10.  
 GN REPLJ OR TTE2304.  
 OS Thermotogaobacter tengcongensis.  
 OC Bacteria; Firmicutes; Clostridia; Thermotogaobacteriales;  
 OC Thermotogaobacteriaceae; Thermotogaobacter.  
 OX NCBI\_TaxID=119072;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=MB4 / JCM 110007;  
 RC MEDLINE=21992816; PubMed=11997336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 Tan H., Chen R., Wang J., Yu J., Yang H.;  
 "A complete sequence of T. tengcongensis genome.";  
 RL Genome Res. 12:689-700(2002).  
 CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AE013173; AAM25445.1; -;  
 DR HAMAP; MF 00362; -; 1.  
 DR InterPro; IPR001790; Ribosomal\_L10.  
 DR InterPro; IPR002363; Ribosomal\_L10eub.  
 DR Pfam; PF00466; Ribosomal\_L10; 1.  
 DR PROSITE; PS01109; RIBOSOMAL\_L10; FALSE NEG.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 177 AA; 19591 MW; 50DDCF896EF6F4E8 CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 177;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 12 VAEF 15  
 -----  
 RESULT 39  
 DSBB\_PASMU  
 ID DSBB\_PASMU STANDARD; PRT; 178 AA.  
 AC Q9L6B3; P57804;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Disulfide bond formation protein B (Disulfide oxidoreductase).  
 GN DSBB OR PM0046.  
 OS Pasteurella multocida.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Pasteurella.  
 OX NCBI\_TaxID=747;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP Fuller T.E., Kennedy M.J., Lowery D.E.;  
 "Identification of Pasteurella multocida virulence genes in a  
 septicemic mouse model using signature-tagged mutagenesis.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Pm70;  
 RC MEDLINE=211145866; PubMed=11248100;  
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
 "Complete genomic sequence of Pasteurella multocida Pm70.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
 RL





TRANSFAC; T01644; -.  
 FlyBase; FBgn0002631; HLFms.  
 GO; GO:0005634; C:nucleus; IDA.  
 GO; GO:0003677; F:DNA binding; IDA.  
 InterPro; IPR001092; HLH\_basic.  
 InterPro; IPR003650; Orange.  
 Pfam; PF00010; HLH; 1.  
 SMART; SM00353; HLH; 1.  
 SMART; SM00511; ORANGE; 1.  
 PROSITE; PS00888; HLH; 1.  
 Differentiation; Neurogenesis; Nuclear protein; DNA-binding;  
 Transcription regulation; Repressor.  
 DNA\_BIND 19 33 BASIC DOMAIN.  
 DOMAIN 34 74 HELIX-LOOP-HELIX MOTIF.  
 DOMAIN 88 129  
 DOMAIN 175 178 WRPW MOTIF.  
 SEQUENCE 178 AA; 19923 MW; 19363D0F6043C84F CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 45 VAEF 48

RESULT 41  
 18 ARCFU STANDARD; PRT; 178 AA.  
 O29147;  
 18-OCT-2001 (Rel. 40, Created)  
 18-OCT-2001 (Rel. 40, Last sequence update)  
 18-OCT-2001 (Rel. 40, Last annotation update)  
 Hypothetical protein AF1118 precursor.  
 AF1118.  
 Archaeoglobus fulgidus.  
 Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 Archaeoglobaceae; Archaeoglobus.  
 NCBI\_TaxID=2234;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 MDLINE=98049343; PubMed:9389475;  
 Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
 Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,  
 Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
 Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
 Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
 Overbeek R., Gocayne J.D., Weidman J.P., McDonald L., Utterback T.,  
 Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,  
 Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
 Mason T.W., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 Venter J.C.;  
 "The complete genome sequence of the hyperthermophilic, sulphate-  
 reducing archaeon Archaeoglobus fulgidus";  
 Nature 390:364-370(1997).  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AE001027; AB090141.1; -.  
 PIR; E69389; E69389.  
 TIGR; AF1118; -.  
 Hypothetical protein; Signal; Complete proteome.  
 SIGNAL 1 20 POTENTIAL.  
 CHAIN 21 178 HYPOTHETICAL PROTEIN AF1118.

SQL SEQUENCE 178 AA; 19319 MW; 378A4F200240D924 CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 155 VAEF 158

RESULT 42  
 APT\_HAEDU STANDARD; PRT; 179 AA.  
 AC QVKK04;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).  
 GN APT OR HD1818  
 OS Haemophilus ducreyi.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OC NCBI\_TaxID=730;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=35000HP / ATCC 700724;  
 RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,  
 Johnson L., Nguyen D., Wang J., Forst C., Hood L.;  
 "The complete genome sequence of Haemophilus ducreyi";  
 Submitted (JUN-2003) to the EMBL/GenBank/DBSJ databases.  
 RL  
 CC -!- FUNCTION: Catalyzes a salvage reaction resulting in the formation  
 of AMP, that is energetically less costly than de novo synthesis.  
 CC -!- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-alpha-  
 D-ribose 1-diphosphate.  
 CC -!- PATHWAY: Purine salvage.  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to the purine/pyrimidine  
 phosphoribosyltransferase family.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; A3017156; AAP96568.1; -.  
 DR HAMAP; MF 00004; -.  
 DR InterPro; IPR005764; Ade phspho trans.  
 DR InterPro; IPR002375; Pr/PY rp transf.  
 DR InterPro; IPR000836; PRTransferase.  
 DR Pfam; PF00156; Pribosyltran; 1.  
 DR TIGRFAMs; TIGR01090; apt; 1.  
 DR PROSITE; PS00103; PUR\_PYR\_PR\_TRANSFER; 1.  
 DR TRANSFERASE; Glycosyltransferase; Purine salvage; Complete proteome.  
 SQL SEQUENCE 179 AA; 19315 MW; 13A1DAFDC13560DF CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 179;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 45 VAEF 48

RESULT 43  
 RETB\_BOVIN STANDARD; PRT; 183 AA.  
 ID RETB\_BOVIN  
 AC P18902;



01-NOV-1990 (Rel. 16, Created)  
01-NOV-1990 (Rel. 16, Last sequence update)  
28-FEB-2003 (Rel. 41, Last annotation update)  
Plasma retinol-binding protein (PRBP) (RBP).  
RBP4.  
Bos taurus (Bovine).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Theria; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
NCBI\_TaxID=9913;  
[1]  
SEQUENCE.  
MEDLINE=91006139; PubMed=2209607;  
Berni R., Stoppani M., Zapponi M.C., Meloni M.L., Monaco H.L.,  
Zanotti G.;  
"The bovine plasma retinol-binding protein. Amino acid sequence,  
interaction with transthyretin, crystallization and preliminary X-ray  
data.";  
Eur. J. Biochem. 192:507-513 (1990).  
[2]  
SEQUENCE OF 52-183 FROM N.A.  
MEDLINE=93385352; PubMed=8373966;  
Liu K.H., Dore J.J., Roberts M.P., Krishnan R., Hopkins F.M.,  
Godkin J.D.;  
"Expression and cellular localization of retinol-binding protein  
messenger ribonucleic acid in bovine blastocysts and extraembryonic  
membranes.";  
Biol. Reprod. 49:393-400 (1993).  
[3]  
X-RAY CRYSTALLOGRAPHY.  
MEDLINE=92322903; PubMed=1623143;  
Monaco H.L., Zanotti G.;  
"Three-dimensional structure and active site of three hydrophobic  
molecule-binding proteins with significant amino acid sequence  
similarity.";  
Biopolymers 32:457-465 (1992).  
[4]  
X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
MEDLINE=93266508; PubMed=8496140;  
Zanotti G., Berni R., Monaco H.L.;  
"Crystal structure of liganded and unliganded forms of bovine plasma  
retinol-binding protein.";  
J. Biol. Chem. 268:10728-10738 (1993).  
-!- FUNCTION: Delivers retinol from the liver stores to the peripheral  
tissues. In plasma, the RBP-retinol complex interacts with  
transthyretin, this prevents its loss by filtration through the  
kidney glomeruli.  
-!- SUBCELLULAR LOCATION: Secreted.  
-!- SIMILARITY: Belongs to the lipocalin family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; S65585; AAB28336.1; --  
PIR; I46955; I46955.  
PIR; S13186; S13186.  
PDB; 1ERB; 31-JAN-94.  
PDB; 1HBP; 31-JAN-94.  
PDB; 1HEQ; 31-JAN-94.  
PDB; 1FEL; 01-NOV-94.  
PDB; 1FEM; 01-NOV-94.  
PDB; 1FEN; 01-NOV-94.  
InterPro; IPR002345; Lipocalin.  
InterPro; IPR000566; Lipocalin\_cytRBP.  
Pfam; PF00061; Lipocalin; 1.  
PRINTS; PR00179; LIPOCALIN.  
PROSITE; PS00213; LIPOCALIN; 1.  
Plasma; Vitamin A; Retinol-binding; Transport; Liver; Lipocalin;

3D-structure. 4 160 BY SIMILARITY.  
FT DISULFID 70 174 BY SIMILARITY.  
FT DISULFID 120 129 BY SIMILARITY.  
FT HELIX 6 8  
FT TURN 13 14  
FT TURN 17 20  
FT STRAND 22 30  
FT TURN 33 34  
FT STRAND 39 47  
FT TURN 49 50  
FT TURN 53 62  
FT TURN 64 65  
FT STRAND 68 79  
FT TURN 83 84  
FT STRAND 85 92  
FT TURN 95 96  
FT STRAND 100 109  
FT STRAND 114 123  
FT TURN 125 126  
FT STRAND 129 138  
FT TURN 141 142  
FT HELIX 146 158  
FT TURN 159 160  
FT TURN 162 163  
FT STRAND 166 167  
FT TURN 173 175  
SQ SEQUENCE 183 AA; 21068 MW; D6BA064CB9E67C09 CRC64;  
Query Match 95.0%; Score 19; DB 1; Length 183;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 42 VAEF 45  
RESULT 44  
DNAA\_WOLSP  
ID DNAA\_WOLSP STANDARD; PRT; 186 AA.  
AC P35907;  
DT 01-JUN-1994 (Rel..29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Chromosomal replication initiator protein dnaa (Fragment).  
GN DNAA.  
OS Wolbachia sp.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
OC Rickettsiaceae; Wolbachieae; Wolbachia.  
OX NCBI\_TaxID=956;  
RN [1]  
SEQUENCE FROM N.A.  
RA MEDLINE=95202115; PubMed=7894745;  
R Bourzais K., Nirgianaki A., Onyango P., Savakis C.;  
RT "A prokaryotic dnaa sequence in Drosophila melanogaster: Wolbachia  
infection and cytoplasmic incompatibility among laboratory strains.";  
RL Insect Mol. Biol. 3:131-142 (1994).  
CC -!- FUNCTION: Plays an important role in the initiation and regulation  
of chromosomal replication. Binds to the origin of replication; it  
binds specifically double-stranded DNA at a 9 bp consensus (dnaa  
box): 5'-TTATC(C/A)A(C/A)A-3'. Dnaa binds to ATP and to acidic  
phospholipids (By similarity).  
CC -!- SIMILARITY: Belongs to the dnaa family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC

```

DR EMBL; Z28981; CAA82285.1; -.
DR PIR; S39317; S39317.
DR HAMAP; MF_00377; -.
DR InterPro; IPR001957; Bac DnaA.
DR Pfam; PF00308; bac_dnaA; 1.
DR PRINTS; PRO0051; DNaA.
DR PROSITE; PS01008; DNaA; PARTIAL.
KW DNA replication; DNA-binding; ATP-binding.
FT NON_TER 1
FT TER 186
SQ SEQUENCE 186 AA; 21283 MW; C423C06CEFFB0459 CRC64;

Query Match          95.0%; Score 19; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 134 VAEF 137

RESULT 45
UCRI_PARDE STANDARD; PRT; 190 AA.
ID UCRI_PARDE
AC P05417;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ubiquinol-cytochrome c reductase iron-sulfur subunit (EC 1.10.2.2)
DE (Rieske iron-sulfur protein) (RISP).
DE PETA.
GN Paracoccus denitrificans.
OS Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8007612; PubMed=2820981;
RA Kurowski B., Ludwig B.;
RT "The genes of the Paracoccus denitrificans bc1 complex. Nucleotide
RT sequence and homologies between bacterial and mitochondrial
RT subunits."
RL J. Biol. Chem. 262:13805-13811(1987).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis.
CC -!- CATALYTIC ACTIVITY: QH(2) + 2 ferricytochrome c = Q + 2
CC ferrocycytochrome c.
CC -!- COFACTOR: Binds 1 2Fe-2S iron sulfur cluster per subunit (By
CC similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein.
CC -!- SUBCELLULAR LOCATION: Membrane-bound.
CC -!- MISCELLANEOUS: The Rieske protein is a high potential 2Fe-2S
CC protein.
CC -!- SIMILARITY: Belongs to the Rieske family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17522; AAA25571.1; -.
CC PIR; A29413; A29413.
CC HSSP; P13272; IRI5.
CC InterPro; IPR005805; Rieske.
CC InterPro; IPR005806; Rieske dom.
CC InterPro; IPR006317; Rieske_proteo.

```

```

DR InterPro; IPR006311; Tat.
DR Pfam; PF00355; Rieske; 1.
DR PRINTS; PRO0162; RIESKE.
DR TIGRFAMs; TIGR01416; Rieske_proteo; 1.
DR TIGRFAMs; TIGR01409; Tat_signal_seq; 1.
DR PROSITE; PS00199; RIESKE_1; 1.
DR PROSITE; PS00200; RIESKE_2; 1.
KW Electron transport; Inner membrane; Transmembrane; Metal-binding;
KW Iron-sulfur; Iron; 2Fe-2S; Oxidoreductase.
FT TRANSMEM 18
FT METAL 132
FT METAL 134
FT METAL 152
FT METAL 155
FT METAL 154
FT DISULFID 137
SQ SEQUENCE 190 AA; 20299 MW; F83F5D9A9C1FEE84 CRC64;

Query Match          95.0%; Score 19; DB 1; Length 190;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 178 VAEF 181

RESULT 46
NTPA_METJA STANDARD; PRT; 193 AA.
ID NTPA_METJA
AC Q57679;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nucleoside-triphosphatase (EC 3.6.1.15) (Nucleoside triphosphate
DE phosphohydrolase) (NTPase).
DE M0226.
GN Methanococcus jannaschii.
OS Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback L.R., Kelliey J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=99332061; PubMed=1040428;
RA Hwang K.Y., Chung J.H., Kim S.-H., Han Y.S., Cho Y.;
RT "Structure-based identification of a novel NTPase from Methanococcus
RT jannaschii."
RL Nat. Struct. Biol. 6:691-696(1999).
CC -!- FUNCTION: CAN EFFICIENTLY HYDROLYZE NONSTANDARD NUCLEOTIDES SUCH
CC AS XTP TO XMP OR ITP TO IMP, BUT NOT THE STANDARD NUCLEOTIDES. XTP
CC IS THE BEST SUBSTRATE.
CC -!- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + phosphate.
CC -!- COFACTOR: Magnesium or manganese.
CC -!- SUBUNIT: Homodimer.
CC -!- SIMILARITY: Belongs to the HAM1 NTPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U67478; BAB98211.1; -  
 PIR; C64328; C64328.  
 PDB; 1B78; 28-JAN-00.  
 PDB; 2MCP; 28-JAN-00.  
 TIGR; MJ0226; -  
 HAMAP; MF\_01405; atypical; 1.  
 InterPro; IPR002637; Hampl like.  
 Pfam; PF01725; Hampl like; 1.  
 TIGRFAMs; TIGR000042; TIGR000042; 1.  
 Hydrolase; Magnesium; 3D-structure; Complete proteome.  
 STRAND 11 14  
 HELIX 18 27  
 TURN 28 30  
 TURN 32 33  
 STRAND 36 39  
 STRAND 46 46  
 HELIX 50 65  
 STRAND 69 78  
 HELIX 79 81  
 TURN 82 83  
 STRAND 85 86  
 TURN 87 88  
 HELIX 89 95  
 TURN 96 96  
 TURN 97 105  
 STRAND 106 107  
 STRAND 112 123  
 TURN 124 125  
 STRAND 126 138  
 HELIX 150 153  
 STRAND 155 157  
 TURN 158 159  
 HELIX 164 166  
 TURN 169 172  
 TURN 173 175  
 HELIX 177 190  
 TURN 191 191  
 SEQUENCE 193 AA; 22202 MW; 3570565E007D3DAC CRC64;

Query Match 95.0%; Score 19; DB 1; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 DB 53 VAEF 56

## RESULT 47

HAM2\_BACHD  
 ID HAM2\_BACHD STANDARD; PRT; 194 AA.  
 AC Q9K8D9;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE HAM1 protein homolog 2.  
 GN BH3067.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirano C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;

"Complete genome sequence of the alkaliphilic bacterium Bacillus RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
 RL Nucleic Acids Res. 28:4317-4331(2000)  
 CC -!- SIMILARITY: Belongs to the HAM1 NTPase family.

-----  
 This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF001517; BAB06786.1; -  
 PIR; C84033; C84033.  
 HSSP; Q57675; 1B78.  
 HAMAP; MF\_01405; -; 1.  
 InterPro; IPR002637; Hampl like.  
 Pfam; PF01725; Hampl like; 1.  
 ProDom; PD004952; Hampl like; 1.  
 TIGRFAMs; TIGR00042; TIGR00042; 1.  
 KW Hydrolase; Complete proteome.  
 SEQUENCE 194 AA; 21868 MW; BCSA0C4F19A04B3 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 DB 14 VAEF 17

## RESULT 48

TNR0\_ECOLI  
 ID TNR0\_ECOLI STANDARD; PRT; 194 AA.  
 AC P05823;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transposon Tn2501 resolvase.  
 GN TNPR.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87109049; PubMed=3027041;  
 RA Michiels T., Cornelis G., Ellis K., Grinstead J.;  
 RT "Tn2501, a component of the lactose transposon Tn951, is an example of a new category of class II transposable elements.";  
 RL J. Bacteriol. 169:624-631(1987).

CC -!- FUNCTION: RESOLVASE CATALYZES THE RESOLUTION (A SITE-SPECIFIC RECOMBINATION) OF THE COINTEGRATED REPLICON TO YIELD THE FINAL TRANSPOSITION PRODUCTS.  
 CC -!- MISCELLANEOUS: Tn2501 IS A CRYPTIC CLASS II TRANSPOSON FOUND AS PART OF THE LACTOSE TRANSPOSON Tn951.  
 CC -!- SIMILARITY: Belongs to the site-specific recombinase resolvase family.

-----  
 This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M15197; AAA27426.1; -  
 PIR; B27758; RPECRS.  
 HSSP; P03012; 2RSL.  
 InterPro; IPR006120; HTH\_7.

```

DR EcGene; Egl1824; pinQ.
DR InterPro; IPR006120; HTH_7.
DR InterPro; IPR006118; Recombinase.
DR InterPro; IPR006119; resolvase_N.
DR Pfam; PF02796; HTH_7; 1.
DR Pfam; PF02939; resolvase; 1.
DR PROSITE; PS00397; RECOMBINASES_1; FALSE_NEG.
DR PROSITE; PS00398; RECOMBINASES_2; 1.
DR Hypothetical protein; DNA recombination; DNA integration; DNA-binding;
KW DNA invertase; Complete proteome.
FT ACT_SITE 11 STRANSIENT COVALENT LINKAGE TO DNA DURING
FT ACT_SITE 11 STRAND CLEAVAGE AND REJOINING
FT FT (BY SIMILARITY).
SQ SEQUENCE 196 AA; 21852 MW; 7CF47F6D8DBD497 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 196;
Best Local Similarity 100.0%; Pred.No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 2 VAEF 5
DB 120 VAEF 123

RESULT 50
PINR_ECOLI
ID ID PINR_ECOLI STANDARD; PRT; 196 AA.
AC P77574;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative DNA-invertase from lambdoid prophage Rac.
GN PINR OR B1374 OR C3146.
DE OS Escherichia coli, and
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TAXID=562, 217992;
[1]
SEQUENCE FROM N.A.
STRAIN=X12 / MG1655;
MEDLINE=97426617; PubMed=9278503;
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12." ;
RL Science 277:1453-1474(1997).
[2]
SEQUENCE FROM N.A.
STRAIN=X12;
MEDLINE=97251357; PubMed=9097039;
Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
Makino K., Mikii T., Mizobuchi K., Mori H., Mori T., Motomura K.,
Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
Sampaio G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
Takenoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-Kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map." ;
RL DNA Res. 3:363-377(1996).
[3]
SEQUENCE FROM N.A.
STRAIN=06:HI / CFT073 / ATCC 700928;
MEDLINE=22388234; PubMed=12471157;
Welch R.A., Burland V., Plunkett G. III, Redford P., Rosesch P.,
Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli." ;
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
[4]
SIMILARITY: Belongs to the site-specific recombinase resolvase

```

```

CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; AE000234; AAC74456.1; -.
CC ENBL; D90775; BAA14979.1; -.
CC ENBL; AE016764; AAN81596.1; -.
CC HSP; P03012; 2RSL.
CC Ecogene; EG13372; pinR.
CC InterPro; IPR006120; HTH_7.
CC InterPro; IPR006118; Recombinase.
CC InterPro; IPR006119; resolvase_N.
CC Pfam; PF02796; HTH_7; 1.
CC Pfam; PF02796; resolvase; 1.
CC PROSITE; PS00397; RECOMBINASES_1; FALSE_NEG.
CC PROSITE; PS00398; RECOMBINASES_2; 1.
CC Hypothetical protein; DNA recombination; DNA integration; DNA-binding;
KW DNA invertase; Complete proteome.
FT ACT_SITE 11 TRANSIENT COVALENT LINKAGE TO DNA DURING
FT ACT_SITE 11 STRAND CLEAVAGE AND REJOINING
FT FT (BY SIMILARITY).
SQ SEQUENCE 196 AA; 21908 MW; 643A62BC940B29A9 CRC64;
Query Match 95.0%; Score 19; DB 1; Length 196;
Basic Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 VASF 5
Db 120 VASF 123
|||||
|||||

RESULT 51
CHPA_ECOLI STANDARD; PRT; 196 AA.
AC P10940;
AT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcriptional regulatory protein uhpA.
GN UHPA OR B3669 OR C4593 OR Z5159 OR ECS4606 OR SP3792 OR S3976.
OS Escherichia coli,
OS Escherichia coli O6,
OS Escherichia coli O157:H7, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 83334, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=87279903; PubMed=3301805;
RA Friedrich M.J., Kadner R.J.;
RT "Nucleotide sequence of the uhp region of Escherichia coli.";
RL J. Bacteriol. 169:3556-3563 (1987).
CC [2]
CC SEQUENCE FROM N.A.
CC SPECIES=E.coli;
CC MEDLINE=92234930; PubMed=1569007;
CC Island M.D., Wei B.-Y., Kadner R.J.;
CC "Structure and function of the uhp genes for the sugar phosphate
CC transport system in Escherichia coli and Salmonella typhimurium.";
CC J. Bacteriol. 174:2754-2762 (1992).
CC [3]
CC SEQUENCE FROM N.A.
CC SPECIES=E.coli; STRAIN=K12 / MG1655;

```

```

RX MEDLINE=93315143; PubMed=7686882;
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
RT genome: organizational symmetry around the origin of replication.";
RN Genomics 16:551-561 (1993).
RN [4]
RN SEQUENCE FROM N.A.
RP SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.P., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
RN [5]
RN SEQUENCE FROM N.A.
RP SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.P., Evans P.S., Gregor J., Kirkpatrick L.A.,
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN [6]
RN SEQUENCE FROM N.A.
RP SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Tida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
RN [7]
RN SEQUENCE FROM N.A.
RP SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441 (2002).
RN [8]
RN SEQUENCE FROM N.A.
RP SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.P., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786 (2003).
CC [9]
CC FUNCTION: APPEARS TO BE THE POSITIVE ACTIVATOR OF UHPT
CC TRANSCRIPTION.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- PTM: PHOSPHORYLATED BY UHPB (PROBABLE).
CC -1- SIMILARITY: Contains 1 response regulatory domain.
CC -1- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```



Davidson N.O., Honjo T.;  
 "Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells.";  
 J. Biol. Chem. 274:18470-18476(1999).  
 C -1- FUNCTION: RNA-editing deaminase involved in somatic hypermutation, gene conversion, and class-switch recombination. Required for several crucial steps of B-cell terminal differentiation necessary for efficient antibody responses.  
 C -1- CATALYTIC ACTIVITY: Cytidine + H(2)O = uridine + NH(3).  
 C -1- COFACTOR: Zinc (By similarity).  
 C -1- SIMILARITY: Belongs to the cytidine and deoxycytidylate deaminases family.  
 C This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 C EMBL: AF132979; AAD41793.1; -.  
 C MGD; MG1342279; Aicda.  
 C GO; GO:0004126; F:cytidine deaminase activity; IDA.  
 C InterPro; IPR007904; APOBEC\_C.  
 C Pfam; PF05240; APOBEC\_C; 1.  
 C PROSITE; PS00903; CYT-DCMP\_DEAMINASES; 1.  
 C METAL 56 56 ZINC (BY SIMILARITY).  
 C METAL 87 87 ZINC (BY SIMILARITY).  
 C METAL 90 90 ZINC (BY SIMILARITY).  
 C SEQUENCE 198 AA; 24030 MW; 18A3BA10CA54BBE2 CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Y 2 VAEF 5  
 D 94 VAEF 97  
 RESULT 54  
 HIS\_SULSO STANDARD; PRT; 199 AA.  
 ID HIS\_SULSO  
 AC O33777;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Imidazole glycerol phosphate synthase subunit high (EC 2.4.2.-) (IGP synthase subunit high) (IMGP synthase subunit high) (IGPS subunit high).  
 DE HIGH (IMGP synthase subunit high) (IGPS subunit high).  
 DE HIGH OR SSC0600 OR C08.050.  
 CS Sulfolobus solfataricus.  
 CC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 CC Sulfolobus.  
 CX NCBI\_taxid=2287;  
 [1]  
 SEQUENCE FROM N.A.  
 RP STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=97352708; PubMed=9209067;  
 RA Charlebois R.L., Sensen C.W., Doolittle W.F., Brown J.R.;  
 RT "Evolutionary analysis of the hisGABDFDEHI gene cluster from the archaeon Sulfolobus solfataricus P2.";  
 RL J. Bacteriol. 179:4429-4432(1997).  
 [2]  
 SEQUENCE FROM N.A.  
 RP STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=20165948; PubMed=10701121;  
 RA Charlebois R.L., Singh R.K., Chan-Weher C.C.-Y., Allard G., Chow C., Confalonieri F., Curtis B., Duguet M., Erauso G., Faguy D.,

Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C., Kushwaha N., Lafleur E., Medina N., Peng X., Penny S.L., She Q., St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F., Ragan M.A., Sensen C.W.;  
 "Gene content and organization of a 281-kbp contig from the genome of the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";  
 Genome 43:116-136(2000).  
 [3]  
 SEQUENCE FROM N.A.  
 RP STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=21332296; PubMed=11427726;  
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G., Aways M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A., De Moers A., Braus G., Fletcher C., Gordon P.M.K., Medina N., Peng X., Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N., Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T., Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;  
 "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";  
 Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
 CC -1- FUNCTION: IGPS catalyzes the conversion of PRAR and glutamine to amidotransferase activity. The high subunit provides the ammonia necessary to hisf for the synthesis of IGP and AICAR (By similarity).  
 CC -1- CATALYTIC ACTIVITY: 5-[(5-phospho-1-deoxyribulose-1-ylamino)methyl]ideneamino-1-(5-phosphoribosyl)imidazole-4-carboxamide + L-glutamine = imidazole-glycerol phosphate + 5-aminoimidazole-4-carboxamide ribonucleotide + L-glutamate + H(2)O.  
 CC -1- PATHWAY: Histidine biosynthesis; fifth step.  
 CC -1- SUBUNIT: Heterodimer of high and hisf (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.  
 C This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 C EMBL: U82227; AAB63025.1; -.  
 C EMBL: Y18930; CAB57700.1; -.  
 C EMBL: AE006689; AAK40911.1; -.  
 C PIR; H90206; H90206.  
 C HAMAP; MF 00278; -. 1.  
 C InterPro; IPR000991; GATase\_1.  
 C Pfam; PF00117; GATase\_1.  
 C PROSITE; PS00442; GATASE\_TYPE I; FALSE NEG.  
 C Histidine biosynthesis; Transferase; Glutamine amidotransferase;  
 C Complete proteome.  
 C ACT\_SITE 76 76 BY SIMILARITY.  
 C ACT\_SITE 178 178 BY SIMILARITY.  
 C ACT\_SITE 180 180 BY SIMILARITY.  
 C SEQUENCE 199 AA; 22533 MW; 17DB1A86724CE095 CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 199;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VAEF 5  
 D 51 VAEF 54  
 RESULT 55  
 RETB\_HORSE STANDARD; PRT; 201 AA.  
 ID RETB\_HORSE  
 AC Q28369;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Plasma retinol-binding protein precursor (PRBP) (RBP).



RP4.  
Equus caballus (Horse).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
NCBI\_TaxID=9796;  
[1]  
SEQUENCE FROM N.A.  
TISSUE=Endometrium;  
MEDLINE=95226610; PubMed=7536053;  
McDowell K.J., Adams M.H., Franklin K.M., Baker C.B.;  
"Changes in equine endometrial retinol-binding protein RNA during the  
estrous cycle and early pregnancy and with exogenous steroids.";  
Biol. Reprod. 52:438-443(1995)  
-1- FUNCTION: Delivers retinol from the liver stores to the peripheral  
tissues. In plasma, the RBP-retinol complex interacts with  
transferrin, this prevents its loss by filtration through the  
kidney glomeruli.  
-1- SUBCELLULAR LOCATION: Secreted.  
-1- SIMILARITY: Belongs to the lipocalin family.  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL: U21208; AAC48461.1; -  
PIR: I45257; I46257.  
HSP: P27485; I4QB.  
InterPro: IPR002345; Lipocalin.  
InterPro: IPR000566; Lipocalin\_cytfabp.  
Pfam: PF00061; lipocalin; 1.  
PRINTS: PR00179; LIPOCALIN.  
PROSITE: PS00213; LIPOCALIN; 1.  
Plasma; Vitamin A; Retinol-binding; Transport; Liver; Signal;  
Lipocalin.  
SIGNAL: 1 18  
CHAIN 19 201  
DISULFID 22 178  
BY SIMILARITY.  
DISULFID 88 192  
BY SIMILARITY.  
DISULFID 138 147  
BY SIMILARITY.  
SEQUENCE 201 AA; 23022 MW; 12CF80834B4262DC CRC64;  
Query Match 95.0%; Score 19; DB 1; Length 201;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 VAEF 5  
|||  
60 VAEF 63  
SULT 56  
TB HUMAN  
RETE HUMAN STANDARD; PRT; 201 AA.  
P02753; O43478; O43479; Q8W3A3; Q9P178;  
21-JUL-1986 (Rel. 01, Created)  
28-FEB-2003 (Rel. 41, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
Plasma retinol-binding protein precursor (PRBP) (RBP) (PRO22222).  
RP4.  
Homo sapiens (Human).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=84069802; PubMed=6316270;  
Colantuoni V., Romano V., Seisi G., Santoro C., Costanzo F.,  
Raugei G., Cortese R.;  
"Cloning and sequencing of a full length cDNA coding for human

retinol-binding protein.";  
Nucleic Acids Res. 11:7769-7776(1983).  
[2]  
SEQUENCE FROM N.A.  
TISSUE=Liver;  
MEDLINE=22388257; PubMed=12477932;  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Sims S.R., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Klausner R.D., Collins F.S., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Altschul S.F., Zeeberg B., Moore T., Max S.J., Wang J., Heish F.,  
Hopkins R.F., Jordan H., Moore T., Max S.J., Wang J., Heish F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Vallaloon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[3]  
SEQUENCE OF 1-189 FROM N.A.  
MEDLINE=86055755; PubMed=2998779;  
D'Onofrio C., Colantuoni V., Cortese R.;  
"Structure and cell-specific expression of a cloned human retinol  
binding protein gene: the 5'-flanking region contains hepatoma  
specific transcriptional signals.";  
EMBO J. 4:1981-1989(1985).  
[4]  
SEQUENCE OF 19-201, AND DISULFIDE BONDS.  
MEDLINE=88019004; PubMed=2444024;  
Rask L., Anundi H., Fohlman J., Peterson P.A.;  
"The complete amino acid sequence of human serum retinol-binding  
protein.";  
Ups. J. Med. Sci. 92:115-146(1987).  
[5]  
SEQUENCE OF 19-201.  
MEDLINE=81254137; PubMed=6942701;  
Rask L., Anundi H., Boehme J., Eriksson U., Ronne H., Sege K.,  
Peterson P.A.;  
"Structural and functional studies of vitamin A-binding proteins.";  
Ann. N.Y. Acad. Sci. 359:79-90(1981).  
[6]  
SEQUENCE OF 19-183.  
MEDLINE=80004132; PubMed=573217;  
Rask L., Anundi H., Peterson P.A.;  
"The primary structure of the human retinol-binding protein.";  
FEBS Lett. 104:55-58(1979).  
[7]  
SEQUENCE OF 18-201 FROM N.A.  
TISSUE=Fetal liver;  
Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,  
He F.;  
"Functional prediction of the coding sequences of 79 new genes deduced  
by analysis of cDNA clones from human fetal liver.";  
Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
[8]  
PARTIAL SEQUENCE.  
MEDLINE=95395382; PubMed=7666002;  
Jaconi S., Rose K., Hughes G.J., Saurat J.-H., Siegenthaler G.;  
"Characterization of two post-translationally processed forms of  
human serum retinol-binding protein: altered ratios in chronic renal  
failure.";  
J. Lipid Res. 36:1247-1253(1995).  
[9]  
X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
MEDLINE=91017498; PubMed=2217163;  
RX



Cowan S.W., Newcomer M.E., Jones T.A.;  
"Crystallographic refinement of human serum retinol binding protein  
at 2-A resolution.";  
Proteins 8:44-61(1990).  
[10]  
X-RAY CRYSTALLOGRAPHY.  
MEDLINE=92322903; PubMed=1623143;  
Monaco H.L., Zanotti G.;  
"Three-dimensional structure and active site of three hydrophobic  
molecule-binding proteins with significant amino acid sequence  
similarity.";  
Biopolymers 32:457-465(1992).  
[11]  
X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS) OF COMPLEX WITH TTR.  
MEDLINE=99162254; PubMed=10052934;  
Naylor H.M., Newcomer M.E.;  
"The structure of human retinol-binding protein (RBP) with its  
carrier protein transthyretin reveals an interaction with the carboxy  
terminus of RBP.";  
Biochemistry 38:2647-2653(1999).  
[12]  
VARIANTS RBP DEFICIENCY ASN-59 AND ASP-93.  
MEDLINE=99103495; PubMed=9888420;  
Seeliger M.W., Biesalski H.K., Wissinger B., Gollnick H., Gielen S.,  
Frank J., Beck S., Zrenner E.;  
"Phenotype in retinol deficiency due to a hereditary defect in retinol  
binding protein synthesis.";  
Invest. Ophthalmol. Vis. Sci. 40:3-11(1999).  
[13]  
CHARACTERIZATION OF VARIANTS ASN-59 AND ASP-93.  
MEDLINE=99247593; PubMed=10232633;  
Biesalski H.K., Frank J., Beck S.C., Heinrich F., Illek B., Reifen R.,  
Gollnick H., Seeliger M.W., Wissinger B., Zrenner E.;  
"Biochemical but not clinical vitamin A deficiency results from  
mutations in the gene for retinol binding protein.";  
Am. J. Clin. Nutr. 69:931-936(1999).  
-!- FUNCTION: Delivers retinol from the liver stores to the peripheral  
tissues. In plasma, the RBP-retinol complex interacts with  
transthyretin, this prevents its loss by filtration through the  
kidney glomeruli.  
-!- SUBCELLULAR LOCATION: Secreted.  
-!- MASS SPECTROMETRY: MW=21063.46; MW\_ERR=1.88; METHOD=Electrospray;  
RANGE=17-199.  
-!- DISEASE: Defects in RBP4 are a cause of retinol-binding protein  
deficiency [MIM:180250]. This condition causes night vision  
problems. It produces a typical "fundus xerophthalmicus."  
-!- DISEASE: A deficiency of vitamin A blocks secretion of the binding  
protein posttranslationally and results in defective delivery and  
supply of vitamin to the epidermal cells (a condition associated  
with a dermatosis).  
-!- SIMILARITY: Belongs to the lipocalin family.  
-!- DATABASE: NAME=Mutations of the RBP4 gene;  
NOTE=Retina International's Scientific Newsletter;  
WWW="http://www.retina-international.com/sci-news/rbp4mut.htm".  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; X00129; CAA24959.1; -;  
EMBL; BC020633; AAH20633.1; -;  
EMBL; X02775; CAA25553.1; -;  
EMBL; X02824; CAB46489.1; -;  
EMBL; AF119868; AAF69622.1; ALT\_INIT.  
EMBL; AF025334; AAC02945.1; -;  
EMBL; AF025335; AAC02946.1; -;  
PIR; A93494; VAHU.  
PDB; 1RBP; 15-JUL-91.

PDB; 1BRP; 31-JAN-94.  
PDB; 1BRQ; 31-JAN-94.  
PDB; 1QAB; 09-APR-99.  
DR SWISS-2DPAGE; P02753; HUMAN.  
DR Siena-2DPAGE; P02753; -;  
DR Genew; HGNC:9922; RBP4.  
DR MIM; 180250; -;  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR GO; GO:0005501; F:retinoid binding; TAS.  
DR InterPro; IPR002345; Lipocalin.  
DR Pfam; PF00061; Lipocalin; 1.  
DR PRINTS; PR00179; LIPOCALIN.  
DR PROSITE; PS00213; LIPOCALIN; 1.  
KW Plasma; Vitamin A; Retinol-binding; Transport; Liver; Signal;  
Lipocalin; Disease mutation; Vision; 3D-structure.  
FT SIGNAL 1 18 PLASMA RETINOL-BINDING PROTEIN.  
FT CHAIN 19 201  
FT DISULFID 22 178  
FT DISULFID 88 192  
FT DISULFID 138 147  
FT VARIANT 59 59  
FT VARIANT 93 93  
FT CONFLICT 8 8 I -> N (in RBP deficiency).  
FT CONFLICT 13 17 /FTid=VAR\_009276.  
FT STRAND 23 23 G -> D (in RBP deficiency).  
FT TURN 31 32 /FTid=VAR\_009277.  
FT STRAND 40 48 F -> L (IN REF. 1 AND 3).  
FT STRAND 60 66 LGSR -> WAA (IN REF. 1 AND 3).  
FT TURN 67 69  
FT STRAND 70 80  
FT TURN 82 83  
FT STRAND 86 96  
Query Match 95.0%; Score 19; DB 1; Length 201;  
Best Local Similarity 100.0%; Pred.No. 3.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 VAEF 5  
DB 60 VAEF 63  
RESULT 57  
ID RETB\_PIG STANDARD; PRT; 201 AA.  
AC P27485;  
DT 01-AUG-1992 (Rel. 23; Created)  
DT 15-JUL-1998 (Rel. 36; Last sequence update)  
DT 28-FEB-2003 (Rel. 41; Last annotation update)  
DE Plasma retinol-binding protein precursor (PRBP) (RBP).  
GN RBP4.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=921331004; PubMed=1723146;  
RA Trout W.E., McDonnell J.J., Kramer K.K., Baumbach G.A., Roberts R.M.;  
RT "The retinol-binding protein of the expanding pig blastocyst:  
disc.";  
RT Mol. Endocrinol. 5:1533-1540(1991).  
RL [2]  
RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS), AND REVISIONS TO 134 AND 185.  
RX MEDLINE=98437649; PubMed=9757135;  
RA Zanotti G., Panzavolta M., Marcato A., Malpeli G., Folli C., Berni R.;  
RT "Structure of pig plasma retinol-binding protein at 1.65-A

resolution.";  
 Acta Crystallogr. D 54:1049-1052(1998).  
 -i- FUNCTION: Delivers retinol from the liver stores to the peripheral tissues. In plasma, the RBP-retinol complex interacts with transthyretin, this prevents its loss by filtration through the kidney glomeruli.  
 -i- SUBCELLULAR LOCATION: Secreted.  
 -i- SIMILARITY: Belongs to the lipocalin family.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; M68860; AAA31113.1; --  
 PIR; A39486; A39486.  
 PDB; 1AQB; 28-JAN-98.  
 InterPro; IPR002345; Lipocalin.  
 InterPro; IPR000566; Lipocalin\_cytFABP.  
 Pfam; PF00061; Lipocalin.1.  
 PRINTS; PR00179; LIPOCALIN.  
 PROSITE; PS00213; LIPOCALIN; 1.  
 Plasma; Vitamin A; Retinol-binding; Transport; Liver; Signal;  
 Lipocalin; 3D-structure.  
 -----  
 SIGNAL 1 18 POTENTIAL.  
 CHAIN 19 201 PLASMA RETINOL-BINDING PROTEIN.  
 DISULFID 22 178  
 DISULFID 88 192  
 DISULFID 138 147  
 CONFLICT 134 134 V -> A (IN REF. 1).  
 CONFLICT 185 185 I -> L (IN REF. 1).  
 STRAND 23 23  
 HELIX 24 26  
 TURN 31 32  
 TURN 35 38  
 HELIX 40 48  
 STRAND 55 65  
 TURN 67 68  
 STRAND 71 80  
 TURN 82 83  
 STRAND 86 97  
 TURN 101 102  
 STRAND 103 110  
 TURN 113 114  
 STRAND 118 127  
 STRAND 132 141  
 TURN 143 144  
 STRAND 146 156  
 TURN 159 160  
 HELIX 164 176  
 TURN 177 178  
 STRAND 180 181  
 TURN 184 185  
 STRAND 191 192  
 TURN 191 192  
 SEQUENCE 201 AA; 23067 MW; A20E39D3C9471DC8 CRC64;  
 -----  
 Query Match 95.0%; Score 19; DB 1; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 ||||  
 60 VAEF 63  
 -----  
 SULT 58  
 TB\_RABIT STANDARD; PRT; 201 AA.  
 P06912;  
 01-JAN-1988 (Rel. 06, Created)

01-OCT-1996 (Rel. 34, Last sequence update)  
 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Plasma retinol-binding protein precursor (PRBP) (RBP).  
 RBP4  
 GN Oryctolagus cuniculus (Rabbit)  
 OS Oryctolagus cuniculus (Rabbit)  
 OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93011736; PubMed=1339354;  
 RA Lee S.Y., Uebels J.L., Soprano D.R.;  
 RT "The lacrimal gland synthesizes retinol-binding protein.";  
 RL Exp. Eye Res. 55:163-171(1992).  
 RN [2]  
 RP SEQUENCE OF 19-201.  
 RX MEDLINE=85207643; PubMed=3838985;  
 RA Sundelin J., Laurent B.C., Anundi H., Traegsaard L., Larhammar D.,  
 RA Bjorck L., Eriksson U., Akerstrom B., Jones A., Newcomer M.,  
 RA Peterson P.A., Rask L.;  
 RT "Amino acid sequence homologies between rabbit, rat, and human serum  
 retinol-binding proteins.";  
 RL J. Biol. Chem. 260:6472-6480(1985).  
 CC -i- FUNCTION: Delivers retinol from the liver stores to the peripheral  
 tissues. In plasma, the RBP-retinol complex interacts with  
 transthyretin, this prevents its loss by filtration through the  
 kidney glomeruli.  
 CC -i- SUBCELLULAR LOCATION: Secreted.  
 CC -i- SIMILARITY: Belongs to the lipocalin family.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; S45958; AAB23582.1; --  
 PIR; A49178; VABP.  
 HSP; P18902; IFEN.  
 DR InterPro; IPR002345; Lipocalin.  
 DR InterPro; IPR000566; Lipocalin\_cytFABP.  
 Pfam; PF00061; Lipocalin.1.  
 PRINTS; PR00179; LIPOCALIN.  
 PROSITE; PS00213; LIPOCALIN; 1.  
 Plasma; Vitamin A; Retinol-binding; Transport; Liver; Lipocalin;  
 Signal.  
 KM SIGNAL 1 18  
 FT CHAIN 19 201 PLASMA RETINOL-BINDING PROTEIN.  
 FT DISULFID 22 178 BY SIMILARITY.  
 FT DISULFID 88 192 BY SIMILARITY.  
 FT DISULFID 138 147 BY SIMILARITY.  
 SQ SEQUENCE 201 AA; 23102 MW; 4153FCF050184136 CRC64;  
 -----  
 Query Match 95.0%; Score 19; DB 1; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 ||||  
 60 VAEF 63  
 -----  
 RESULT 59  
 YIGZ\_ECOLI STANDARD; PRT; 204 AA.  
 ID YIGZ\_ECOLI  
 AC P27852;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein yigZ.  
 GN YIGZ OR B3848.



```

CC -1- SUBUNIT: INTERACTS WITH CDC20 AND WITH ADAM17 (TACE).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE MAD2 FAMILY.
CC -1- SIMILARITY: Contains 1 HORMA domain.
CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/MAD2L1ID304.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; U85410; AAC50781.1; -.
DR ENBL; U31278; AAC52060.1; -.
DR ENBL; AJ000186; CAA03943.1; -.
DR ENBL; AB056160; BAB63410.1; -.
DR ENBL; BC000356; AAH00356.1; -.
DR ENBL; BC005945; AAH05945.1; -.
DR PIR; G01942; G01942.
DR PDB; 1DUJ; 08-MAR-00.
DR Genew; HGNC:6763; MAD2L1.
DR GK; Q13257; -.
DR MIM; 601467; -.
DR GO; GO:0005699; C:kinetochore; TAS.
DR GO; GO:0007067; P:mitosis; TAS.
DR GO; GO:0007093; P:mitotic checkpoint; TAS.
DR InterPro; IPR003511; DNABind_HORMA.
DR Pfam; PF02301; HORMA; 1.
DR PROSITE; PS50815; HORMA; 1.
KW Cell cycle; Mitosis; Nuclear protein; 3D-structure.
DR DOMAIN 14 197
FT STRAND 11 12
FT STRAND 15 15
FT HELIX 17 34
FT TURN 35 35
FT TURN 41 42
FT STRAND 44 48
FT TURN 49 50
FT STRAND 51 55
FT HELIX 59 73
FT STRAND 83 90
FT TURN 91 94
FT STRAND 95 103
FT STRAND 106 106
FT TURN 112 113
FT HELIX 122 141
FT STRAND 149 158
FT TURN 165 166
FT STRAND 173 175
FT TURN 179 181
FT STRAND 185 193
SQ SEQUENCE 205 AA; 23510 MW; B8DCBF0043836764 CRC64;
Query Match 95.0%; Score 19; DB 1; Length 205;
Best Local Similarity 100.0%; Pred.No.3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 VAEF 5
Db 20 VAEF 23
RESULT 61
MD21_MOUSE
ID MD21_MOUSE STANDARD; PRT; 205 AA.
AC Q921B5; Q9J153;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mitotic spindle assembly checkpoint protein MAD2A (MAD2-like 1).

```

```

GN MAD2L1 OR MAD2A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RA Jin D.-Y.; Jeang K.-T.;
RT "Identification of a novel component of the spindle assembly
RT checkpoint in mammalian cells.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=20348739; PubMed=10892650;
RA Dobles M.; Liberal V.; Scott M.L.; Beneser R.; Sorger P.K.;
RT "Chromosome missegregation and apoptosis in mice lacking the mitotic
RT checkpoint protein Mad2.";
RL Cell 101:635-645(2000).
CC -1- FUNCTION: REQUIRED FOR THE EXECUTION OF THE MITOTIC CHECKPOINT
CC WHICH MONITORS THE PROCESS OF KINETOCORE-SPINDLE ATTACHMENT AND
CC DELAYS THE ONSET OF ANAPHASE WHEN THIS PROCESS IS NOT COMPLETE. IT
CC INHIBITS THE ACTIVITY OF THE ANAPHASE PROMOTING COMPLEX BY
CC SEQUESTERING CDC20 UNTIL ALL CHROMOSOMES ARE ALIGNED AT THE
CC METAPHASE PLATE (BY SIMILARITY).
CC -1- SUBUNIT: INTERACTS WITH CDC20.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE MAD2 FAMILY.
CC -1- SIMILARITY: Contains 1 HORMA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; U83902; AAD09238.1; -.
DR ENBL; AF261919; AAF69525.1; -.
DR HSSP; Q13257; 1DUJ.
DR MGD; MGI:1860374; Mad2l1.
DR GO; GO:0000776; C:kinetochore; IDA.
DR GO; GO:0005816; C:spindle pole body; IDA.
DR GO; GO:0000707; P:mitotic chromosome segregation; IMP.
DR GO; GO:0007094; P:mitotic spindle checkpoint; IMP.
DR InterPro; IPR003511; DNABind_HORMA.
DR Pfam; PF02301; HORMA; 1.
DR PROSITE; PS50815; HORMA; 1.
KW Cell cycle; Mitosis; Nuclear protein.
FT DOMAIN 14 197
FT CONFLICT 157 157 T -> A (IN REF. 1).
FT CONFLICT 178 178 C -> S (IN REF. 1).
FT CONFLICT 201 201 T -> I (IN REF. 1).
SQ SEQUENCE 205 AA; 23598 MW; A9F3E28BC4C9738E CRC64;
Query Match 95.0%; Score 19; DB 1; Length 205;
Best Local Similarity 100.0%; Pred.No.3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 VAEF 5
Db 20 VAEF 23
RESULT 62
YAC1_LEGPN
ID YAC1_LEGPN STANDARD; PRT; 208 AA.
AC P37033;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Hypothetical 23.7 kDa protein in acn 5'/region.
OS Legionella pneumophila.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=446;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Philadelphia 1;
RX MEDLINE=93374864; PubMed=8366052;
RA Mengaud J.M., Horwitz M.A.;
RT "The major iron-containing protein of Legionella pneumophila is an
RT aconitase homologous with the human iron-responsive element-binding
RT protein.";
RL J. Bacteriol. 175:5666-5676(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L22081; AAA25294.1; -.
DR PIR; A48642; A48642.
KW Hypothetical protein.
SQ SEQUENCE. 208 AA; 23714 MW; 155C75R63A20C800 CRC64;

Query Match          95.0%; Score 19; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 45 VAEF 48

RESULT 63
HMOO_CORDI
ID HMOO_CORDI STANDARD; PRT; 215 AA.
AC P71119;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Heme oxygenase (EC 1.14.99.3).
GN HMOO OR DP1669.
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SC7;
RX MEDLINE=97158681; PubMed=9006041;
RA Schmitt M.P.;
RT "Utilization of host iron sources by Corynebacterium diphtheriae:
RT identification of a gene whose product is homologous to eukaryotic
RT heme oxygenases and is required for acquisition of iron from heme and
RT hemoglobin.";
RL J. Bacteriol. 179:838-845(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Biotype Gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910;
RX Cerdeno-Taraga A.M., Efstathiou A., Dover L.G., Holden M.T.G.,
RX Pallen M., Bentley S.D., Besra G.S., Churcher C., James K.D.,
RX De Goya A., Childingworth T., Cronin A., Dowd L., Feltwell T.,
RX Hamlin N., Holroyd S., Jagsis K., Moule S., Quail M.A.,
RX Rabinowitsch E., Rutherford K.M., Thomson N.R., Urwin L.,
RX Whitehead S., Barrell B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
RT diphtheriae NCTC13129."
RL Nucleic Acids Res. 31:6516-6523(2003).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U73860; AAC44832.1; -.
DR EMBL; BX248358; CAE50198.1; -.
DR HSP; P09601; IQQ8.
DR InterPro; IPR002051; Heme_oxygenase.
DR Pfam; PF01126; Heme_oxygenase; 1.
DR PRINTS; PR00088; HAEMOXIGNASE.
DR PROSITE; PS00593; HEME_OXYGENASE; 1.
KW Heme; Oxidoreductase.
FT METAL 20 20 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT CONFLICT 34 34 E -> K (IN REF. 1).
FT CONFLICT 60 60 A -> V (IN REF. 1).
FT CONFLICT 92 93 DG -> GS (IN REF. 1).
FT CONFLICT 192 192 N -> H (IN REF. 1).
SQ SEQUENCE 215 AA; 24116 MW; 60D9E8E2ED7ED456 CRC64;

Query Match          95.0%; Score 19; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 39 VAEF 42

RESULT 64
COAT_CWIX
ID COAT_CWIX STANDARD; PRT; 218 AA.
AC Q66120;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Coat protein (Capsid protein) (CP).
OS Cucumber mosaic virus (strain Ixora) (CMV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Cucumovirus.
OX NCBI_TaxID=117114;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96005047; PubMed=7561763;
RA McCarvey P.B., Foushant M., Geleka L., Cellini F., Kaper J.M.;
RT "The complete sequence of a cucumber mosaic virus from Ixora that is
RT deficient in the replication of satellite RNAs.";
RL J. Gen. Virol. 76:2257-2270(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U20219; AAC54619.1; -.
DR PIR; D71392; D71392.
DR InterPro; IPR000247; Cucumovirus_coat.
DR InterPro; IPR008975; Viral_cap_coat.

```

PFam; PF00760; Cucumo\_coat; 1.  
 PRINTS; PR00222; CUCUMOCOAT.  
 ProDom; PD001284; Cucumovirus\_coat; 1.  
 Coat protein.  
 SEQUENCE 218 AA; 24185 MW; 130582D17BD75224 CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 218;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 |||||  
 96 VAEF 99  
 RESULT 66  
 CRI\_ARCFU STANDARD; PRT; 223 AA.  
 D ECK1\_ARCFU  
 C Q29758;  
 T 10-OCT-2003 (Rel. 42, Created)

PTP1\_NPVOP STANDARD; PRT; 220 AA.  
 O10274;  
 01-NOV-1997 (Rel. 35, Created)  
 01-NOV-1997 (Rel. 35, Last sequence update)  
 10-OCT-2003 (Rel. 42, Last annotation update)  
 Putative protein-tyrosine phosphatase 1 (EC 3.1.3.48).  
 PTP-1.  
 Orygia pseudotsugata multicapsid polyhedrosis virus (OpMPV).  
 Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 Nucleopolyhedrovirus.  
 NCBI\_TaxID=164623;  
 [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=9721300; PubMed=9126251;  
 Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,  
 Rohrmann G.F.;  
 "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear  
 polyhedrosis virus genome.";  
 Virology 229:381-393(1997).  
 -!- FUNCTION: COULD BE INACTIVE AS THE ACTIVE SITE CYSTEINE IS  
 MODIFIED TO TRYPTOPHAN.  
 -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein  
 tyrosine + phosphate.  
 -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
 TYROSINE PHOSPHATASE FAMILY. CDC14 SUBFAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; U75930; AAC59009.1; -;  
 InterPro; IPR000387; TYR\_phosphatase.  
 PROSITE; PS00383; TYR\_PHOSPHATASE\_1; FALSE\_NEG.  
 PROSITE; PSS00386; TYR\_PHOSPHATASE\_2; 1.  
 4 Hydrolase.  
 SITE 162 162 EQUIVALENT OF ACTIVE SITE CYS.  
 SEQUENCE 220 AA; 25234 MW; D3FC033F1953D425 CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 |||||  
 139 VAEF 142

DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Probable exosome complex RNA-binding protein 1.  
 GN AF0492  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 OC Archaeoglobaceae; Archaeoglobus.  
 OX NCBI\_TaxID=2234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE=98049343; PubMed=938475;  
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,  
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,  
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 reducing archaeon Archaeoglobus fulgidus.";  
 RL Nature 390:364-370(1997).  
 CC -!- FUNCTION: Probably involved in degradation of a variety of RNA  
 species; could act a RNA-binding component of the exosome  
 (Potential).  
 CC -!- SUBUNIT: Component of the archaeal exosome multi-enzyme  
 ribonuclease complex (Potential).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -!- SIMILARITY: Contains 1 KH domain.  
 CC -!- SIMILARITY: Contains 1 S1 motif domain.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AE001070; AAB90745.1; -;  
 PIR; D69311; D69311.  
 TIGR; AF0492; -;  
 HAMAP; MF\_00623; -; 1.  
 InterPro; IPR004087; KH\_dom.  
 InterPro; IPR004088; KH\_type\_1.  
 InterPro; IPR008994; Nucleic\_acid\_OB.  
 InterPro; IPR001029; S1.  
 Pfam; PF00013; KH; 1.  
 Pfam; PF00575; S1; 1.  
 SMART; SM00322; KH; 1.  
 SMART; SM00316; S1; 1.  
 PROSITE; PSS0084; KH\_TYPE\_1; 1.  
 PROSITE; PSS0126; S1; 1.  
 KW Exosome; RNA-binding; Complete proteome.  
 FT DOMAIN 58 127 KH  
 FT MOTIF.  
 SEQUENCE 223 AA; 24725 MW; A1B7F1A26972833 CRC64;  
 Query Match 95.0%; Score 19; DB 1; Length 223;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 VAEF 5  
 |||||  
 209 VAEF 212  
 RESULT 67  
 RAN\_TETPY

```

D  RAN TETPY STANDARD; PRT; 223 AA.
C  P41914;
T  01-NOV-1995 (Rel. 32, Created)
T  01-NOV-1995 (Rel. 32, Last sequence update)
T  28-FEB-2003 (Rel. 41, Last annotation update)
S  GTP-binding nuclear protein RAN/TC4.
E  Tetrahymena pyriformis.
C  Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
C  Tetrahymena; Tetrahymena.
C  NCBI_TaxID=5908;
X  [1]
N  SEQUENCE FROM N.A.
P  STRAIN=W;
C  MEDLINE=94299154; PubMed=8026746;
C  Nagata K.-I., Takemasa T., Alam S., Hattori T., Watanabe Y.,
A  Nozawa Y.;
T  "Cloning of cDNAs encoding a cell-cycle-regulatory GTP-binding
T  low-M(r) (GBLM) protein, Ran/TC4, from micronucleated Tetrahymena
T  thermophila and amiconucleated Tetrahymena pyriformis.";
C  Gene 144:123-125 (1994).
L  -1- FUNCTION: GTP-binding protein involved in nucleocytoplasmic
C  transport. Required for the import of protein into the nucleus and
C  also for RNA export. Involved in chromatin condensation and
C  control of cell cycle (By similarity).
C  -1- SUBUNIT: Monomer.
C  -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
C  -1- SIMILARITY: Belongs to the small GTPase superfamily. Ran family.
C  -----
C  This SWISS-PROT entry is copyright. It is produced through a collaboration
C  between the Swiss Institute of Bioinformatics and the EMBL outstation -
C  the European Bioinformatics Institute. There are no restrictions on its
C  use by non-profit institutions as long as its content is in no way
C  modified and this statement is not removed. Usage by and for commercial
C  entities requires a license agreement (See http://www.isb-sib.ch/announce/
C  or send an email to license@isb-sib.ch).
C  -----
C  EMBL: D21825; BAA04849.1; --
C  HSSP: P17080; 1A2K.
R  InterPro: IPR002041; RAN.
R  InterPro: IPR001806; Ras_trnsmfmg.
R  InterPro: IPR005225; Small_GTP.
R  Pfam: PF00071; ras; 1.
R  PRINTS: PR00627; GTPRANTC4.
R  PRINTS: PR00449; RASTRNSFRMG.
R  SMART: SM00176; RAN; 1.
R  TIGRFAMs: TIGR00233; small_GTP; 1.
R  PROSITE: PS01115; RAN; 1.
C  GTP-binding; Nuclear protein; Protein transport.
C  NP_BIND 18 25 GTP (BY SIMILARITY).
C  NP_BIND 66 70 GTP (BY SIMILARITY).
C  NP_BIND 123 126 GTP (BY SIMILARITY).
C  DOMAIN 128 143 IBB (BY SIMILARITY).
C  SEQUENCE 223 AA; 25422 MW; 5474DAD58AF0977 CRC64;
S  Query Match 95.0%; Score 19; DB 1; Length 223;
S  Best Local Similarity 100.0%; Pred. No. 3.9e+02;
S  Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2Y 2 VAEP 5
2b 9 VAEP 12
RESULT 68
VATD TETH STANDARD; PRT; 223 AA.
AC 087860;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE V-type ATP synthase subunit D (EC 3.6.3.14) (V-type ATPase subunit D).
DE ATPD OR VATD.
DE Thermus thermophilus.
CS

```

```

OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OC NCBI_TaxID=274;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=HB8 / ATCC 27634;
RC MEDLINE=2050364; PubMed=10789522;
RX Yokoyama K., Ohkuma S., Teguchi H., Yasunaga T., Wakabayashi T.,
RA Yoshida M.;
RT "V-type H+-ATPase/synthase from a thermophilic eubacterium, Thermus
RT thermophilus. Subunit structure and operon.";
RL J. Biol. Chem. 275:13955-13961 (2000).
CC -1- FUNCTION: Produces ATP from ADP in the presence of a proton
CC gradient across the membrane.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -1- SIMILARITY: Belongs to the V-ATPase D subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D63799; BAA33198.2; --
CC HAMAP: MF_00271; 1.
DR InterPro: IPR002699; ATPsynth_Deub.
DR Pfam: PF01813; ATP-synt_D; 1.
DR ProDom: PD004122; ATPsynth_Deub; 1.
DR TIGRFAMs: TIGR00309; V_ATPase_subd; 1.
KW Hydrolase; ATP synthetase; Hydrogen ion transport.
SQ SEQUENCE 223 AA; 24677 MW; B666AB00DF0B08B CRC64;
S  Query Match 95.0%; Score 19; DB 1; Length 223;
S  Best Local Similarity 100.0%; Pred. No. 3.9e+02;
S  Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 VAEP 5
Db 36 VAEP 39
RESULT 69
RAN TETH STANDARD; PRT; 225 AA.
AC P41915;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE GTP-binding nuclear protein RAN/TC4.
DE Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymena; Tetrahymena.
OC NCBI_TaxID=5911;
X  [1]
N  SEQUENCE FROM N.A.
P  STRAIN=B1868;
C  MEDLINE=94299154; PubMed=8026746;
C  Nagata K.-I., Takemasa T., Alam S., Hattori T., Watanabe Y.,
RA Nozawa Y.;
RT "Cloning of cDNAs encoding a cell-cycle-regulatory GTP-binding
RT low-M(r) (GBLM) protein, Ran/TC4, from micronucleated Tetrahymena
RT thermophila and amiconucleated Tetrahymena pyriformis.";
C  Gene 144:123-125 (1994).
L  -1- FUNCTION: GTP-binding protein involved in nucleocytoplasmic
C  transport. Required for the import of protein into the nucleus and
C  also for RNA export. Involved in chromatin condensation and
C  control of cell cycle (By similarity).
C  -1- SUBUNIT: Monomer.
C  -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
C  -1- SIMILARITY: Belongs to the small GTPase superfamily. Ran family.
CC

```

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; D17748; BAA04600.1; -;  
 HSP; P17080; LA2K.  
 InterPro; IPR002041; RAN.  
 InterPro; IPR001806; Ras trnsfrmg.  
 InterPro; IPR005225; Small\_GTP.  
 Pfam; PF00071; ras; 1.  
 PRINTS; PR00627; GTPRANTC4.  
 PRINTS; PR00449; RASTRNSPRMG.  
 SMART; SM00176; RAN; 1.  
 TIGRFAMs; TIGR00231; small\_GTP; 1.  
 PROSITE; PS01115; RAN; 1.  
 GTP-binding; Nuclear protein; Protein transport.  
 NP BIND 18 25 GTP (BY SIMILARITY).  
 NP BIND 66 70 GTP (BY SIMILARITY).  
 NP BIND 123 126 GTP (BY SIMILARITY).  
 DOMAIN 128 143 IBB (BY SIMILARITY).  
 SEQUENCE 225 AA; 25649 MW; 69F8B8409AF703F8 CRC64;

Query Match 95.0%; Score 19; DB 1; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02; Indels 0;  
 Matches 4; Conservative 0; Mismatches 0; Gaps 0;

2 VAEF 5  
 ||||  
 9 VAEF 12

SULT 70  
 32 NPVOP

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; D13796; BAA02952.1; -;  
 EMBL; D13923; BAA03030.1; -;  
 EMBL; U75930; AAC59129.1; -;  
 DR PIR; D30857; D30857.  
 KW Hypothetical protein; Late protein.  
 SQ SEQUENCE 228 AA; 25962 MW; F0766AD21F60039D CRC64;

Query Match 95.0%; Score 19; DB 1; Length 228;

Best Local Similarity 100.0%; Pred. No. 4e+02; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0; Gaps 0;

QY 2 VAEF 5  
 ||||  
 DB 84 VAEF 87

RESULT 71

NADD\_PSEPK STANDARD; PRT; 230 AA.  
 AC Q88DL5;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Probable nicotinate-nucleotide adenyltransferase (EC 2.7.7.18)  
 DE (Deamido-NAD(+) pyrophosphorylase) (Deamido-NAD(+) diphosphorylase)  
 DE (Nicotinate mononucleotide adenyltransferase) (NAmN  
 DE adenyltransferase).  
 DE NADD OR PP4810.  
 GN Pseudomonas putida (strain KT2440).  
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=160488;  
 RN [1]

SEQUENCE FROM N.A.  
 MEDLINE=22423060; PubMed=12534463;  
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
 RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,  
 RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,  
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,  
 RA Uterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,  
 RA Lauber J., Stjepandic D., Hobeisel J., Straetz M., Heim S.,  
 RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoft A., Tsemmler B.,  
 RA Fraser C.M.;  
 RA "Complete genome sequence and comparative analysis of the  
 RA metabolically versatile *Pseudomonas putida* KT2440.";  
 RA Environ. Microbiol. 4:799-808(2002).  
 RL ENVIRON. MICROBIOL. 4:799-808(2002).  
 CC -!- FUNCTION: Catalyzes the reversible adenylation of nicotinate  
 CC mononucleotide (NAmN) to nicotinic acid adenine dinucleotide  
 CC (NAD) (By similarity).  
 CC -!- CATALYTIC ACTIVITY: ATP + nicotinate ribonucleotide = diphosphate  
 CC + deamido-NAD(+).  
 CC -!- PATHWAY: NAD biosynthesis.  
 CC -!- SIMILARITY: Belongs to the nadD family.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; AE016792; AAN70379.1; -;  
 DR TIGR; PF4810; -;  
 DR HAVAP; MF\_00244; -; 1.  
 DR InterPro; IPR004820; Cytidylyltransf.  
 DR Pfam; PF01467; CTP transf 2; 1.  
 KW Pyridine nucleotide biosynthesis; Transferase; Nucleotidyltransferase;  
 KW NAD; Complete proteome.  
 SQ SEQUENCE 230 AA; 25291 MW; 4CF04E8F70B48941 CRC64;



```

Query Match          95.0%; Score 19; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5
b 42 VAEF 45

RESULT 72
R3_PORPU STANDARD; PRT; 230 AA.
D RR3_PORPU
C P51308;
T 01-OCT-1996 (Rel. 34, Created)
T 01-OCT-1996 (Rel. 34, Last sequence update)
T 28-FEB-2003 (Rel. 41, Last annotation update)
E Chloroplast 30S ribosomal protein S3.
N RPS3.
S Porphyra purpurea.
G Chloroplast.
C Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
X NCBI_TaxID=278;
P SEQUENCE FROM N.A.
C STRAIN=Avonport;
A Keith M.E., Munnolland J.;
T "Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.";
L Plant Mol. Biol. Rep. 13:333-335(1995).
C -1- SUBUNIT; Part of the 30S ribosomal subunit.
C -1- SUBCELLULAR LOCATION: Chloroplast.
C -1- SIMILARITY: Belongs to the S3P family of ribosomal proteins.
C -1- SIMILARITY: Contains 1 KH type-2 domain.
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See http://www.isb-sib.ch/announce/
C or send an email to license@isb-sib.ch).
C -----
C EMBL; U38804; AAC08194.1; -.
R PIR; S73229; S73229.
R HAMAP; MF_01309; 1.
R InterPro; IPR004087; KH dom.
R InterPro; IPR009019; KH_prok.
R InterPro; IPR004044; KH_TYPE_2.
R InterPro; IPR001351; Ribosomal_S3_C.
R InterPro; IPR008282; Ribosomal_S3_N.
R InterPro; IPR005704; S3_bact.
R Pfam; PF00013; KH; 1.
R Pfam; PF00189; Ribosomal_S3_C; 1.
R Pfam; PF00417; Ribosomal_S3_N; 1.
R SMART; SM00322; KH; 1.
R TIGRFSMS; TIGR01009; rpsC_bact; 1.
R PROSITE; PS50823; KH_TYPE_2; 1.
R PROSITE; PS00548; RIBOSOMAL_S3; 1.
W Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
T DONALDIN 39 109 KH TYPE-2.
Q SEQUENCE 230 AA; 25631 MW; 5489B93704D7F0ED CRC64;

Query Match          95.0%; Score 19; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5
b 118 VAEF 121

RESULT 73
BIE_COREP

```

```

ID UBIE COREP STANDARD; PRT; 230 AA.
AC Q8PSB3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Menquinone biosynthesis methyltransferase ubiE (EC 2.1.1.-).
GN UBIE OR C60481.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=22723752; PubMed=12840036;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579(2003).
CC -1- FUNCTION: Methyltransferase required for the conversion of
CC dimethylmenaquinone (DMKH2) to menaquinone (MKH2) (By similarity).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =
CC S-adenosyl-L-homocysteine + menaquinol.
CC -1- PATHWAY: Menaquinone biosynthesis; last step.
CC -1- SIMILARITY: Belongs to the ubiE family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF005215; BAC17291.1; ALT_INIT.
DR HAMAP; MF_01813; 1.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR000051; SAM_Bind.
DR InterPro; IPR004033; UbiE/COQ5_Metrf.
DR Pfam; PF01209; UbiE_methyltran; 1.
DR PROSITE; PS01183; UbiE_1; FALSE NEG.
DR PROSITE; PS01184; UbiE_2; FALSE NEG.
KW Menaquinone biosynthesis; Transferase; Methyltransferase;
KW Complete proteome.
KW SEQUENCE 230 AA; 25366 MW; E10DCB602A1CD886 CRC64;

Query Match          95.0%; Score 19; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
Db 145 VAEF 148

RESULT 74
UBIE CORGL STANDARD; PRT; 230 AA.
ID UBIE CORGL
AC Q8NT39;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Menaquinone biosynthesis methyltransferase ubiE (EC 2.1.1.-).
GN UBIE OR CGL0471.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
SEQUENCE FROM N.A.

```

RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
RA Nakagawa S.;  
RT "Complete genomic sequence of *Corynebacterium glutamicum* ATCC 13032.";  
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Methyltransferase required for the conversion of  
CC dimethylmenaquinone (DMK2) to menaquinone (MK2) (By similarity).  
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =  
CC S-adenosyl-L-homocysteine + menaquinol.  
CC -!- PATHWAY: Menaquinone biosynthesis; last step.  
CC -!- SIMILARITY: Belongs to the ubiE family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AP005275; BAB97864.1; -.  
CC HAMAP; MF 01813; -; 1.  
CC InterPro; IPR001601; Methyltransf.  
CC InterPro; IPR000051; SAM bind.  
CC InterPro; IPR004033; UbiE/COQ5 Methyl.  
CC Pfam; PF01209; UbiE\_methyltransf; 1.  
CC PROSITE; PS01183; UBI1\_1; FALSE NEG.  
CC PROSITE; PS01184; UBI2; FALSE NEG.  
CC Menaquinone biosynthesis; Transferase; Methyltransferase;  
CC Complete proteome.  
CC SEQUENCE 230 AA; 25244 MW; A8548173B7901400 CRC64;  
CC -----  
CC Query Match 95.0%; Score 19; DB 1; Length 230;  
CC Best Local Similarity 100.0%; Pred. No. 4e+02;  
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CC -----  
CC 2 VAEF 5  
CC 145 VAEF 148  
CC ||||  
CC -----  
CC RESULT 75  
CC ID VGI8 ECOLI STANDARD; PRT; 234 AA.  
CC P24155;  
CC 01-MAR-1992 (Rel. 21, Created)  
CC 01-MAR-1992 (Rel. 21, Last sequence update)  
CC 10-OCT-2003 (Rel. 42, Last annotation update)  
CC DE Hypothetical protein ygiB.  
CC VGI8 OR B3037 OR C3783 OR SF3077 OR S3282.  
CC S Escherichia coli.  
CC Escherichia coli.  
CC Shigella flexneri.  
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
CC Enterobacteriaceae; Escherichia.  
CC NCBI\_TaxID=562, 217992, 623;  
CC [1]  
CC SEQUENCE FROM N.A.  
CC SPECIES=E. coli;  
CC MEDLINE=92223101; PubMed=1314093;  
CC Yang T.-P., Depew R.S.;  
CC "Nucleotide sequence of a region duplicated in *Escherichia coli* toc  
CC mutants.";  
CC Biochim. Biophys. Acta 1130:227-228(1992).  
CC [2]  
CC SEQUENCE FROM N.A.  
CC SPECIES=E. coli; STRAIN=K12 / MG1655;  
CC MEDLINE=97428617; PubMed=9278503;  
CC Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
CC Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
CC Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
CC Mau B., Shao Y.;  
CC "The complete genome sequence of *Escherichia coli* K-12.";  
CC Science 277:1453-1474(1997).  
CC -----

[3]  
SEQUENCE FROM N.A.  
SPECIES=E. coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;  
MEDLINE=22388234; PubMed=12471157;  
Weich R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
Raasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
"Extensive mosaic structure revealed by the complete genome sequence  
of uropathogenic *Escherichia coli*.";  
Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
[4]  
SEQUENCE FROM N.A.  
SPECIES=S. flexneri; STRAIN=301 / Serotype 2a;  
MEDLINE=2272406; PubMed=12384590;  
Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
Jiang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
Yu J.;  
"Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity  
through comparison with genomes of *Escherichia coli* K12 and O157.";  
Nucleic Acids Res. 30:4432-4441(2002).  
[5]  
SEQUENCE FROM N.A.  
SPECIES=S. flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;  
MEDLINE=22590274; PubMed=12704152;  
Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,  
Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,  
Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,  
Schwartz D.C., Blattner F.R.;  
"Complete genome sequence and comparative genomics of *Shigella*  
*flexneri* serotype 2a strain 2457T.";  
Infect. Immun. 71:2775-2786(2003).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M77129; AAA71875.1; -.  
EMBL; U28377; AAA62205.1; -.  
EMBL; AE000385; AAC76073.1; -.  
EMBL; AE016766; AAN82227.1; -.  
EMBL; AE015319; AAN4555.1; ALT INIT.  
EMBL; AE016988; AAP18368.1; -.  
PIR; S22360; S22360.  
ECOGene; EG11164; ygiB.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 234 AA; 24868 MW; 08D5DE931072C966 CRC64;  
-----  
Query Match 95.0%; Score 19; DB 1; Length 234;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
QY 2 VAEF 5  
DB 100 VAEF 103  
-----  
Search completed: May 24, 2004, 17:41:02  
Job time : 58.5714 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: May 24, 2004, 17:33:08 ; Search time 35.7143 Seconds

(without alignments)

44.173 Million cell updates/sec

US-09-594-978a-1

20

1 XVAEP 5

BLOSUM62

Gapop 10.0 , Gapext 0.5

1017041 seqs, 315518202 residues

1017041

Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

SPTREML 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriapi:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

result NO.	Score	Query Match	Length	ID	Description
1	19	95.0	20	6	Q9T518
2	19	95.0	22	2	Q9R5C0
3	19	95.0	30	5	P82214
4	19	95.0	33	16	C25518
5	19	95.0	35	17	Q8Z219
6	19	95.0	36	6	Q9TQR7
7	19	95.0	44	2	Q48420
8	19	95.0	45	16	Q830H9
9	19	95.0	48	16	Q25478
10	19	95.0	51	16	Q7UXL5
11	19	95.0	52	2	Q44941
12	19	95.0	54	16	Q8X3Y8
13	19	95.0	54	16	Q8FIF3
14	19	95.0	61	2	Q8V8B8
15	19	95.0	61	4	Q9P166
16	19	95.0	68	12	Q99GZ9

17	19	95.0	68	16	Q82X12	Q82xi2 nitrosomona
18	19	95.0	69	10	Q7X7L1	Q7x7l1 oryza sativ
19	19	95.0	69	16	Q8NPA7	Q8npa7 corynebacte
20	19	95.0	71	13	Q8AUQ2	Q8auq2 salmo trutt
21	19	95.0	71	13	Q8AUP0	Q8aup0 salmo salar
22	19	95.0	72	2	P97252	P97252 escherichia
23	19	95.0	74	16	Q97JU4	Q97ju4 clostridium
24	19	95.0	74	16	Q82MX0	Q82mx0 streptomyce
25	19	95.0	77	6	Q862M8	Q862m8 bos taurus
26	19	95.0	77	16	Q88VS7	Q88vs7 lactobacill
27	19	95.0	79	12	Q98582	Q98582 paramecium
28	19	95.0	80	2	Q9XI36	Q9xi36 agrobacteri
29	19	95.0	81	16	Q99WL6	Q99wl6 bradyrhizob
30	19	95.0	84	16	Q9KAV4	Q9kav4 bacillus ha
31	19	95.0	84	16	Q821D2	Q821d2 chlamydophi
32	19	95.0	84	16	Q7VPV6	Q7vpv6 chlamydia p
33	19	95.0	85	3	Q43116	Q43116 colletotric
34	19	95.0	87	6	Q93149	Q93149 ovis aries
35	19	95.0	87	9	Q7X3Y9	Q7x3y9 bacterioph
36	19	95.0	88	2	Q98590	Q98590 nitrosomona
37	19	95.0	88	10	Q41185	Q41185 arabidopsis
38	19	95.0	89	2	Q9ZGY2	Q9zgy2 yersinia pe
39	19	95.0	89	16	Q9CNX1	Q9cnx1 pasteurella
40	19	95.0	89	16	Q9CFD1	Q9cfd1 lactococcus
41	19	95.0	90	4	Q81ZX3	Q81zx3 homo sapien
42	19	95.0	91	13	Q98TB2	Q98tb2 ambloplites
43	19	95.0	94	2	Q8KFI2	Q8kfi2 rhizobium e
44	19	95.0	94	5	Q46425	Q46425 cryotolagus
45	19	95.0	96	16	Q8U8Q1	Q8u8q1 agrobacteri
46	19	95.0	96	16	Q89WX7	Q89wx7 bradyrhizob
47	19	95.0	97	12	Q8U0G9	Q8u0g9 pyrococcus
48	19	95.0	97	12	Q86535	Q86535 hepatitis a
49	19	95.0	97	16	Q8Y476	Q8y476 listeria mo
50	19	95.0	99	16	Q9HZ38	Q9hz38 pseudomonas
51	19	95.0	99	16	Q885M3	Q885m3 pseudomonas
52	19	95.0	100	16	Q8ZUR2	Q8zur2 nitrosomona
53	19	95.0	101	2	Q8VS54	Q8vs54 lactobacill
54	19	95.0	101	17	Q77YC2	Q77yc2 sulfolobus
55	19	95.0	102	16	Q81R27	Q81r27 bacillus an
56	19	95.0	103	4	Q722N2	Q722n2 homo sapien
57	19	95.0	103	16	Q9F368	Q9f368 streptomyce
58	19	95.0	103	16	Q89MS9	Q89ms9 bradyrhizob
59	19	95.0	104	9	Q856B5	Q856b5 mycobacteri
60	19	95.0	105	5	Q8SVK0	Q8svk0 encephalito
61	19	95.0	105	9	Q856U1	Q856u1 mycobacteri
62	19	95.0	105	17	Q96Z31	Q96z31 sulfolobus
63	19	95.0	106	1	P71530	P71530 methanococc
64	19	95.0	106	5	Q9BGL5	Q9bgl5 ovis aries
65	19	95.0	106	10	Q23682	Q23682 arabidopsis
66	19	95.0	107	3	Q9C1L1	Q9c1l1 neurospora
67	19	95.0	107	5	Q81HV3	Q81hv3 plasmodium
68	19	95.0	107	16	Q8NRX6	Q8nrx6 corynebacte
69	19	95.0	109	13	Q9PUB2	Q9pu82 crocodylus
70	19	95.0	109	16	Q53333	Q53333 mycobacteri
71	19	95.0	109	16	Q7TX28	Q7tx28 mycobacteri
72	19	95.0	110	12	Q11343	Q11343 molluscum c
73	19	95.0	110	16	Q88EG3	Q88eg3 pseudomonas
74	19	95.0	111	17	Q8U296	Q8u296 pyrococcus
75	19	95.0	112	5	Q86PL2	Q86pl2 caenorhabdi
76	19	95.0	112	13	Q98SR4	Q98sr4 acipenser s
77	19	95.0	112	13	Q98SR3	Q98sr3 acipenser s
78	19	95.0	112	16	Q91315	Q91315 pseudomonas
79	19	95.0	113	2	P95437	P95437 pseudomonas
80	19	95.0	113	2	C69133	C69133 burkholderi
81	19	95.0	113	17	Q8TYL3	Q8tyl3 methanopyru
82	19	95.0	114	16	Q8U5K8	Q8u5k8 agrobacteri
83	19	95.0	115	4	Q9Y5V1	Q9y5v1 homo sapien
84	19	95.0	115	17	Q9V203	Q9v203 pyrococcus
85	19	95.0	116	11	Q8CD94	Q8cd94 mus musculu
86	19	95.0	116	12	Q9DW67	Q9dw67 rat cytomeg
87	19	95.0	118	4	Q8TCV1	Q8tcv1 homo sapien
88	19	95.0	118	10	Q7X1W4	Q7x1w4 oryza sativ
89	19	95.0	119	2	Q9EXG6	Q9exg6 listeria mo

90	19	95.0	119	5	Q95Y64	Q95Y64 caenorhabdi	163	19	95.0	140	13	Q8JIE5	Q8JIE5 agalynchnis
91	19	95.0	119	5	Q95Y64	Q85J0 mycobacteri	164	19	95.0	140	16	Q89IC6	Q89IC6 clostridium
92	19	95.0	119	16	Q92R99	Q92R99 rhizobium m	165	19	95.0	140	16	Q83LW2	Q83LW2 shigella fl
93	19	95.0	119	16	Q87R53	Q87R53 vibrio para	166	19	95.0	142	10	Q23641	Q23641 arabisopsis
94	19	95.0	120	2	Q9ALD4	Q9ALD4 uncultured	167	19	95.0	142	10	Q39368	Q39368 brassica ol
95	19	95.0	120	2	Q9ALH9	Q9ALH9 uncultured	168	19	95.0	142	10	Q41213	Q41213 brassica na
96	19	95.0	120	2	Q9ALB1	Q9ALB1 uncultured	169	19	95.0	142	10	Q42386	Q42386 brassica na
97	19	95.0	120	2	Q31353	Q31353 bacillus ce	170	19	95.0	143	2	Q4526	Q4526 streptomyc
98	19	95.0	120	2	Q9ALH7	Q9ALH7 uncultured	171	19	95.0	143	16	Q8DKS6	Q8DKS6 synchococc
99	19	95.0	120	2	Q9ALH2	Q9ALH2 uncultured	172	19	95.0	143	16	Q82V53	Q82V53 nitrosomona
100	19	95.0	120	2	Q9ALD3	Q9ALD3 uncultured	173	19	95.0	144	5	Q24728	Q24728 drosophila
101	19	95.0	120	2	Q9ALD3	Q9ALD3 uncultured	174	19	95.0	144	10	Q9FFG7	Q9FFG7 arabidopsis
102	19	95.0	120	16	Q8A661	Q8A661 bacteroides	175	19	95.0	144	16	Q8ABF9	Q8ABF9 bacteroides
103	19	95.0	121	4	Q7Z2N5	Q7Z2N5 homo sapien	176	19	95.0	145	2	Q54808	Q54808 streptomyc
104	19	95.0	121	9	Q9B066	Q9B066 mycobacteri	177	19	95.0	145	2	Q84127	Q84127 vibrio vuln
105	19	95.0	122	0	Q7XJZ7	Q7XJZ7 oryza sativ	178	19	95.0	145	13	Q9D8E4	Q9D8E4 oncorhynch
106	19	95.0	122	16	Q9CHJ9	Q9CHJ9 xyliella fas	179	19	95.0	145	16	Q55965	Q55965 synchocyst
107	19	95.0	123	16	Q9CJH8	Q9CJH8 pasteurilla	180	19	95.0	146	5	Q8WTF1	Q8WTF1 ascaris suu
108	19	95.0	123	16	Q9J2J4	Q9J2J4 streptomyc	181	19	95.0	146	5	Q23842	Q23842 drosophila
109	19	95.0	124	9	Q03976	Q03976 bacterioph	182	19	95.0	146	10	Q94AG7	Q94AG7 arabidopsis
110	19	95.0	124	16	Q8FQZ0	Q8FQZ0 corynebacte	183	19	95.0	146	16	Q98UX5	Q98UX5 rhizobium l
111	19	95.0	125	16	Q818K2	Q818K2 bacillus ce	184	19	95.0	146	16	Q8FEX6	Q8FEX6 escherichia
112	19	95.0	125	16	Q69643	Q69643 mycobacteri	185	19	95.0	147	5	Q23929	Q23929 streptococ
113	19	95.0	125	16	Q7TVZ0	Q7TVZ0 mycobacteri	186	19	95.0	147	5	Q24730	Q24730 drosophila
114	19	95.0	125	17	Q26336	Q26336 methanobact	187	19	95.0	147	16	Q69906	Q69906 streptomyc
115	19	95.0	126	2	Q939V9	Q939V9 anabaena so	188	19	95.0	147	17	Q9HPZ8	Q9HPZ8 halobacteri
116	19	95.0	126	6	Q9X3F8	Q9X3F8 bos taurus	189	19	95.0	148	11	Q9CR18	Q9CR18 mus muscul
117	19	95.0	126	12	Q84351	Q84351 simian viru	190	19	95.0	148	16	Q97SS2	Q97SS2 neisseria m
118	19	95.0	126	16	Q9PNZ4	Q9PNZ4 campylobact	191	19	95.0	149	2	Q93KN2	Q93KN2 uncultured
119	19	95.0	126	16	Q7WFX5	Q7WFX5 bordetella	192	19	95.0	149	16	Q89C75	Q89C75 rhizobium l
120	19	95.0	127	12	Q91BF3	Q91BF3 epodoptera	193	19	95.0	149	16	Q8Y9M8	Q8Y9M8 listeria mo
121	19	95.0	127	16	Q69714	Q69714 mycobacteri	194	19	95.0	149	16	Q8FRU2	Q8FRU2 corynebacte
122	19	95.0	127	16	Q928V1	Q928V1 listeria in	195	19	95.0	149	17	Q8LNU8	Q8LNU8 methanosarc
123	19	95.0	127	16	Q83ET6	Q83ET6 coxiella bu	196	19	95.0	150	16	Q92RPE	Q92RPE listeria in
124	19	95.0	127	16	Q7TVS6	Q7TVS6 mycobacteri	197	19	95.0	150	16	Q8YAK5	Q8YAK5 listeria mo
125	19	95.0	128	2	Q938X8	Q938X8 acinetobact	198	19	95.0	151	11	Q9CVJ5	Q9CVJ5 mus muscul
126	19	95.0	128	2	Q83YU7	Q83YU7 acinetobact	199	19	95.0	151	16	Q9ABU9	Q9ABU9 caulobacter
127	19	95.0	128	10	Q9ZSM5	Q9ZSM5 dendrobium	200	19	95.0	152	9	Q8SBJ6	Q8SBJ6 bacterioph
128	19	95.0	128	16	Q9L5K0	Q9L5K0 salmonella	201	19	95.0	152	9	Q7V5S0	Q7V5S0 bacterioph
129	19	95.0	129	10	Q9SBB1	Q9SBB1 arabidopsis	202	19	95.0	153	10	Q9ASA1	Q9ASA1 oryza sativ
130	19	95.0	129	12	Q84339	Q84339 simian viru	203	19	95.0	153	10	Q8S817	Q8S817 oryza sativ
131	19	95.0	130	9	Q8W5Y1	Q8W5Y1 bacterioph	204	19	95.0	153	16	Q8XR05	Q8XR05 ralstonia s
132	19	95.0	130	12	Q84339	Q84339 simian viru	205	19	95.0	153	16	Q88SK9	Q88SK9 lactobacill
133	19	95.0	131	2	Q84Y7C4	Q84Y7C4 tt virus. o	206	19	95.0	154	2	Q52066	Q52066 pantoce agg
134	19	95.0	132	2	Q84CY0	Q84CY0 gamma-prote	207	19	95.0	154	3	Q862D3	Q862D3 fusarium sp
135	19	95.0	132	9	Q8SDF0	Q8SDF0 anabaena sp	208	19	95.0	155	16	Q7UWB1	Q7UWB1 rhodospirell
136	19	95.0	132	16	Q86418	Q86418 pseudomonas	209	19	95.0	156	16	Q9PMR4	Q9PMR4 campylobact
137	19	95.0	133	2	Q8GH32	Q8GH32 gamma-prote	210	19	95.0	156	17	Q974D5	Q974D5 sulfobus
138	19	95.0	133	16	Q9RZR3	Q9RZR3 deihococcus	211	19	95.0	157	5	Q9GTJ0	Q9GTJ0 heterodera
139	19	95.0	133	16	Q8EGZ5	Q8EGZ5 yersinia pe	212	19	95.0	157	16	Q8FA79	Q8FA79 escherichia
140	19	95.0	133	16	Q8D9V9	Q8D9V9 vibrio vuln	213	19	95.0	157	16	Q8DMU7	Q8DMU7 synchococc
141	19	95.0	133	17	Q8U1Z0	Q8U1Z0 pyrococcus	214	19	95.0	159	12	Q84662	Q84662 paramectum
142	19	95.0	134	2	Q937K7	Q937K7 erwinia chr	215	19	95.0	159	16	Q8FTW6	Q8FTW6 corynebacte
143	19	95.0	134	5	Q96898	Q96898 sacculina c	216	19	95.0	160	2	Q93KL6	Q93KL6 uncultured
144	19	95.0	134	9	Q857W7	Q857W7 mycobacteri	217	19	95.0	161	2	Q99Q13	Q99Q13 shigella fl
145	19	95.0	134	16	Q82Q77	Q82Q77 salmonella	218	19	95.0	161	2	Q52115	Q52115 streptomyc
146	19	95.0	134	16	Q828C2	Q828C2 salmonella	219	19	95.0	161	2	Q8V818	Q8V818 shigella fl
147	19	95.0	134	17	Q8Q0F7	Q8Q0F7 methanosarc	220	19	95.0	161	10	Q9AS45	Q9AS45 oryza sativ
148	19	95.0	134	17	Q8Q0F7	Q8Q0F7 methanosarc	221	19	95.0	161	10	Q7XX54	Q7XX54 oryza sativ
149	19	95.0	135	2	Q939V2	Q939V2 anabaena sp	222	19	95.0	161	16	Q98FZ2	Q98FZ2 rhizobium l
150	19	95.0	135	16	Q986D5	Q986D5 rhizobium l	223	19	95.0	161	16	Q8ZL98	Q8ZL98 salmonella
151	19	95.0	135	17	Q96282	Q96282 sulfobus	224	19	95.0	161	16	Q89G21	Q89G21 bradyrhizob
152	19	95.0	136	16	P74444	P74444 synchocyst	225	19	95.0	161	16	Q8Z2B1	Q8Z2B1 salmonella
153	19	95.0	137	2	Q9XBH6	Q9XBH6 bacillus ce	226	19	95.0	162	4	Q9BS36	Q9BS36 homo sapien
154	19	95.0	137	16	Q9PP94	Q9PP94 campylobact	227	19	95.0	162	11	Q8CG34	Q8CG34 mus sp. myb
155	19	95.0	137	16	Q8IX11	Q8IX11 bacillus an	228	19	95.0	163	4	Q86W13	Q86W13 homo sapien
156	19	95.0	137	16	Q914R9	Q914R9 bacillus ce	229	19	95.0	163	13	Q7ZXF9	Q7ZXF9 xenopus lae
157	19	95.0	138	16	Q9KQ08	Q9KQ08 vibrio chol	230	19	95.0	163	16	Q07237	Q07237 mycobacteri
158	19	95.0	138	16	Q98XN0	Q98XN0 bradyrhizob	231	19	95.0	163	16	Q7U2A7	Q7U2A7 mycobacteri
159	19	95.0	139	16	Q9K9P4	Q9K9P4 bacillus ha	232	19	95.0	163	17	Q8ZX97	Q8ZX97 pyrobaculum
160	19	95.0	139	16	Q8D6C8	Q8D6C8 vibrio vuln	233	19	95.0	164	16	Q8Y7N8	Q8Y7N8 listeria mo
161	19	95.0	139	16	Q82NK0	Q82NK0 streptomyc	234	19	95.0	164	16	Q9KZL0	Q9KZL0 streptomyc
162	19	95.0	140	11	Q9ESY3	Q9ESY3 mus muscul	235	19	95.0	164	17	Q29902	Q29902 archaeoglob

236	19	95.0	165	4	Q96ED1	Q96ed1 homo sapien	309	19	95.0	188	16	Q8XC46	Q8xc46 escherichia
237	19	95.0	165	5	Q8SRD1	Q8srd1 encephalito	310	19	95.0	188	17	Q8TNJ0	Q8tnj0 methanosarc
238	19	95.0	165	11	Q9D6S8	Q9d6s8 mus musculus	311	19	95.0	189	2	Q84CM0	Q84cm0 streptomyce
239	19	95.0	166	2	Q52661	Q52661 rhodobacter	312	19	95.0	189	5	Q20199	Q20199 caenorhabdi
240	19	95.0	166	10	Q8R296	Q8r296 oryza sativ	313	19	95.0	189	8	Q8WJ84	Q8wj84 eriosorus h
241	19	95.0	166	10	Q948R0	Q948r0 oryza sativ	314	19	95.0	189	10	Q07299	Q07299 zea mays (m
242	19	95.0	166	10	Q8ZVJ9	Q8zvj9 pyrobaculum	315	19	95.0	189	16	Q8U549	Q8u549 agrobacteri
243	19	95.0	167	17	Q8ZVJ9	Q8zvj9 pyrobaculum	316	19	95.0	189	16	Q8FC12	Q8fc12 escherichia
244	19	95.0	167	2	Q8GAC0	Q8gac0 arthrobacte	317	19	95.0	189	4	Q96F12	Q96f12 homo sapien
245	19	95.0	168	2	Q47850	Q47850 pantoaea agg	318	19	95.0	190	4	Q93RS6	Q93rs6 streptomyce
246	19	95.0	168	16	Q8ZRF0	Q8zrf0 yersinia pe	319	19	95.0	190	16	Q93RS6	Q93rs6 streptomyce
247	19	95.0	169	16	Q9X8N2	Q9x8n2 streptomyce	320	19	95.0	191	2	Q93KL3	Q93kl3 uncultured
248	19	95.0	169	16	Q7UKV3	Q7ukv3 rhodopirell	321	19	95.0	191	10	Q84LW2	Q84lw2 oryza sativ
249	19	95.0	169	16	Q7USD8	Q7usd8 synechococ	322	19	95.0	191	10	Q84LW2	Q84lw2 oryza sativ
250	19	95.0	170	16	Q93JC6	Q93jc6 streptomyce	323	19	95.0	191	16	Q9KKV1	Q9kkv1 vibrio chol
251	19	95.0	170	16	Q881F4	Q881f4 pseudomonas	324	19	95.0	191	16	Q8EAK2	Q8eak2 shewanella
252	19	95.0	171	10	Q84QL9	Q84ql9 oryza sativ	325	19	95.0	192	10	Q8L647	Q8l647 deschampsia
253	19	95.0	171	12	Q8JQ15	Q8jq15 black queen	326	19	95.0	192	16	Q8UG64	Q8ug64 agrobacteri
254	19	95.0	172	11	Q9AWP5	Q9awp5 oryza sativ	327	19	95.0	192	16	Q9ADP7	Q9adp7 streptomyce
255	19	95.0	172	11	Q55179	Q55179 mus musculu	328	19	95.0	192	16	Q83D37	Q83d37 coxiella bu
256	19	95.0	172	11	Q8K164	Q8k164 mus musculu	329	19	95.0	193	2	F74983	F74983 yersinia en
257	19	95.0	172	11	Q55177	Q55177 mus musculu	330	19	95.0	193	16	Q9KY43	Q9ky43 streptomyce
258	19	95.0	172	17	Q9U236	Q9u236 pyrococcus	331	19	95.0	193	16	Q82B77	Q82b77 streptomyce
259	19	95.0	172	17	Q8U1U0	Q8u1u0 pyrococcus	332	19	95.0	194	2	Q9EVD1	Q9evd1 erwinia amy
260	19	95.0	173	5	Q20526	Q20526 caenorhabdi	333	19	95.0	194	2	Q8GQ21	Q8gq21 serratia ma
261	19	95.0	173	5	Q8SWX9	Q8swx9 drosophila	334	19	95.0	194	2	Q8GE12	Q8ge12 erwinia pyr
262	19	95.0	174	11	Q8CTG2	Q8ctg2 mus musculu	335	19	95.0	194	2	Q83XU2	Q83xj2 vibrio angu
263	19	95.0	174	16	Q8ECL6	Q8ec16 shewanella	336	19	95.0	194	16	Q9A7S1	Q9a7s1 caulobacter
264	19	95.0	174	16	Q33289	Q33289 mycobacteri	337	19	95.0	194	16	Q8UBS8	Q8ubs8 agrobacteri
265	19	95.0	174	16	Q7V6A9	Q7v6a9 prochloroco	338	19	95.0	195	10	Q9LMI9	Q9lmi9 arabisopsia
266	19	95.0	174	16	Q7TX22	Q7tx22 mycobacteri	339	19	95.0	195	16	Q89TJ4	Q89tj4 bradyrhizob
267	19	95.0	175	11	Q925A9	Q925a9 rattus norv	340	19	95.0	195	16	Q88SG8	Q88sg8 lactobacill
268	19	95.0	175	16	Q8A2B2	Q8a2b2 bacteroides	341	19	95.0	196	2	Q93KM6	Q93km6 uncultured
269	19	95.0	175	16	Q82DE1	Q82del1 streptomyce	342	19	95.0	196	5	Q9GNR5	Q9gnr5 stytonychia
270	19	95.0	175	17	Q58465	Q58465 pyrococcus	343	19	95.0	196	5	Q9GNR5	Q9gnr5 stytonychia
271	19	95.0	176	16	Q7VD73	Q7vd73 prochloroco	344	19	95.0	196	5	Q9GNR7	Q9gnr7 stytonychia
272	19	95.0	176	16	Q7VD73	Q7vd73 prochloroco	345	19	95.0	196	5	Q9GNR7	Q9gnr7 stytonychia
273	19	95.0	178	5	Q8MTA3	Q8mta3 drosophila	346	19	95.0	196	11	Q9CVX4	Q9cvx4 mus musculu
274	19	95.0	178	16	Q7VJ04	Q7vj04 helicobacte	347	19	95.0	196	11	Q9FTV5	Q9ftv5 corynebacte
275	19	95.0	178	16	Q7UT14	Q7uti14 rhodopirell	348	19	95.0	196	16	Q8DGL7	Q8dgl7 synechococ
276	19	95.0	179	2	Q54157	Q54157 shigella fl	349	19	95.0	196	16	Q8Z2K7	Q8z2k7 salmonella
277	19	95.0	179	11	Q8R057	Q8r057 mus musculu	350	19	95.0	197	2	Q93KN1	Q93kni uncultured
278	19	95.0	179	16	Q7VXQ4	Q7vxq4 haemophilus	351	19	95.0	197	9	Q8HQ1	Q8hq1 vibrio harv
279	19	95.0	180	2	Q8KU52	Q8ku52 shigella fl	352	19	95.0	197	16	Q9CGV5	Q9cgv5 lactococcus
280	19	95.0	180	2	Q8KU52	Q8ku52 shigella fl	353	19	95.0	197	17	Q8TZV4	Q8tzv4 pyrococcus
281	19	95.0	180	10	Q8RXU1	Q8rxu1 arabidopsis	354	19	95.0	198	2	Q8KAL5	Q8kalc6 rhizobium 1
282	19	95.0	180	10	Q9M098	Q9m098 arabidopsis	355	19	95.0	198	16	Q9BNC6	Q9bnc6 rhizobium 1
283	19	95.0	180	10	Q8LDX3	Q8ldx3 arabidopsis	356	19	95.0	198	16	Q50927	Q50927 nitrosomona
284	19	95.0	180	16	Q8ZPH9	Q8zph9 salmonella	357	19	95.0	199	4	Q9H315	Q9h315 homo sapien
285	19	95.0	180	16	Q8Z707	Q8z707 salmonella	358	19	95.0	199	4	Q86X83	Q86x83 homo sapien
286	19	95.0	180	16	Q83T30	Q83t30 salmonella	359	19	95.0	199	11	Q8BXC6	Q8bxc6 mus musculu
287	19	95.0	180	16	Q82BEC1	Q82b21 streptomyce	360	19	95.0	199	11	Q8BJI3	Q8bjj3 mus musculu
288	19	95.0	180	16	Q81DE6	Q81de6 bacillus ce	361	19	95.0	199	16	Q8XTM5	Q8xtm5 raietonia s
289	19	95.0	180	16	Q05559	Q05559 mycobacteri	362	19	95.0	199	16	Q7U983	Q7u983 synechococ
290	19	95.0	180	16	Q7UAE1	Q7uae1 shigella fl	363	19	95.0	200	16	Q8ZMY2	Q8zmy2 salmonella
291	19	95.0	181	5	Q8MUA9	Q8mu9 ancylostoma	364	19	95.0	200	16	Q8ZQA6	Q8zqa6 salmonella
292	19	95.0	181	16	Q8YBS5	Q8ybs5 anabaena sp	365	19	95.0	200	16	Q8ZTW8	Q8ztw8 salmonella
293	19	95.0	182	16	Q7V2G9	Q7v2g9 prochloroco	366	19	95.0	201	2	Q9X407	Q9x407 methylosulf
294	19	95.0	182	16	Q8TW72	Q8twr2 methanopyru	367	19	95.0	201	2	Q9X407	Q9x407 methylosulf
295	19	95.0	183	4	Q8TCQ9	Q8tcq9 homo sapien	368	19	95.0	201	10	Q48848	Q48848 arabidopsis
296	19	95.0	184	16	Q8X5C1	Q8x5c1 escherichia	369	19	95.0	201	16	Q98127	Q98127 rhizobium 1
297	19	95.0	184	16	Q9ZLY9	Q9zly9 helicobacte	370	19	95.0	202	2	Q8RNS6	Q8rns6 lactobacill
298	19	95.0	184	16	Q8Y7R2	Q8y7r2 pseudomonas	371	19	95.0	202	16	Q8YBY2	Q8yby2 listeria mo
299	19	95.0	184	16	Q7VC97	Q7vc97 prochloroco	372	19	95.0	202	16	Q69982	Q69982 streptomyce
300	19	95.0	185	16	Q8UHUS	Q8uhus agrobacteri	373	19	95.0	202	16	P73372	P73372 synechocyst
301	19	95.0	185	16	Q9RZH9	Q9rzh9 deinococcus	374	19	95.0	203	2	Q9RBB8	Q9rb8 arthrobacte
302	19	95.0	186	2	Q56745	Q56745 wolbachia s	375	19	95.0	203	2	Q9A9A7	Q9a9a7 oryza sativ
303	19	95.0	187	2	P72209	P72209 proteus mir	376	19	95.0	203	16	Q8PB95	Q8pb95 xanthomonas
304	19	95.0	187	2	Q93T67	Q93t67 riemeirella	377	19	95.0	203	16	Q898Y8	Q89y8 xanthomonas
305	19	95.0	187	2	Q98FP5	Q98fd5 rhizobium 1	378	19	95.0	203	16	Q88DQ5	Q88dq5 pseudomonas
306	19	95.0	187	16	Q92W50	Q92wb0 rhizobium m	379	19	95.0	204	2	Q9RBE6	Q9rb6 xanthomonas
307	19	95.0	188	5	Q8T3L7	Q8t3l7 drosophila	380	19	95.0	204	2	Q83XE6	Q83xe6 xanthomonas
308	19	95.0	188	16	Q83722	Q83722 treponema p	381	19	95.0	204	4	Q9NZ83	Q9nz83 homo sapien
	19	95.0	188	16	Q8ZH37	Q8zh37 yersinia pe		19	95.0	204	4	Q7Z2N7	Q7z2n7 homo sapien

382	19	95.0	204	16	Q92P26	Q92fz6 rickettsia	455	19	95.0	224	16	Q9NRP2	Q9nrp2 corynebacte
383	19	95.0	204	16	Q8PQC4	Q8pqc4 xanthomonas	456	19	95.0	225	16	Q986L5	Q986l5 rhizobium l
384	19	95.0	204	16	Q93J78	Q93j78 streptomyce	457	19	95.0	225	16	Q825R9	Q825r9 streptomyce
385	19	95.0	204	16	Q7UB59	Q7ub59 shigella fl	458	19	95.0	226	2	Q9RBD7	Q9rbd7 arthrobacte
386	19	95.0	205	16	Q8X310	Q8x310 escherichia	459	19	95.0	226	2	Q848B9	Q848b9 streptomyce
387	19	95.0	205	16	Q8G5L8	Q8g5l8 bifidobacte	460	19	95.0	226	5	Q8WPD3	Q8wpd3 paracentrot
388	19	95.0	205	16	Q8FB10	Q8fb10 escherichia	461	19	95.0	226	11	Q8CSE1	Q8cse1 mus musculu
389	19	95.0	205	16	Q8DVM4	Q8dvm4 streptococc	462	19	95.0	226	16	Q8YDM2	Q8ydm2 brucella me
390	19	95.0	205	16	Q83IV9	Q83iv9 shigella fl	463	19	95.0	226	16	Q8FUS4	Q8fus4 brucella su
391	19	95.0	206	2	Q9KWD3	Q9kwd3 agrobacteri	464	19	95.0	227	2	Q9KJZ2	Q9kjj2 pseudomonas
392	19	95.0	206	11	Q8COY5	Q8coy5 mus musculu	465	19	95.0	227	10	Q8GRV7	Q8grv7 oryza sativ
393	19	95.0	206	16	Q8EC34	Q8ec34 shewanella	466	19	95.0	227	17	Q8ZVD4	Q8zvd4 pyrobaculum
394	19	95.0	207	10	Q8H4R2	Q8h4r2 oryza sativ	467	19	95.0	228	16	Q8Z6Q3	Q8z6q3 salmonella
395	19	95.0	208	16	Q8PD60	Q8pd60 xanthomonas	468	19	95.0	228	16	Q9RS24	Q9rs24 deinococcus
396	19	95.0	208	16	Q82AM2	Q82am2 streptomyce	469	19	95.0	228	16	Q8Z1S3	Q8z1s3 yersinia pe
397	19	95.0	209	2	Q9RB86	Q9rb86 arthrobacte	470	19	95.0	228	16	Q8FSE2	Q8fse2 corynebacte
398	19	95.0	209	4	Q9BYW9	Q9byw9 homo sapien	471	19	95.0	228	16	Q894X0	Q894x0 clostridium
399	19	95.0	209	16	Q8UP55	Q8uf55 agrobacteri	472	19	95.0	228	16	Q83T36	Q83t36 salmonella
400	19	95.0	210	16	Q8ZPD7	Q8zpd7 salmonella	473	19	95.0	229	10	Q9CASS	Q9cas5 arabidopsis
401	19	95.0	210	16	Q8G4Z4	Q8g4z4 bifidobacte	474	19	95.0	229	16	Q667S3	Q667s3 aquifex ae
402	19	95.0	210	16	Q83L90	Q83l90 shigella fl	475	19	95.0	229	16	Q8CSD8	Q8csd8 staphylococ
403	19	95.0	211	10	Q7XA92	Q7xa92 fragaria an	476	19	95.0	229	16	Q832X6	Q832x6 enterococcu
404	19	95.0	211	16	Q92RS0	Q92rs0 rhizobium m	477	19	95.0	230	2	Q9RBB0	Q9rbbo arthrobacte
405	19	95.0	211	16	Q8P940	Q8p940 xanthomonas	478	19	95.0	230	5	O76405	O76405 caenorhabdi
406	19	95.0	212	2	Q3RBB7	Q3rb7 arthrobacte	479	19	95.0	230	10	Q84RH7	Q84rh7 lycopersico
407	19	95.0	212	10	Q8RUH2	Q8ruh2 oryza sativ	480	19	95.0	230	16	Q8UHL0	Q8uhl0 agrobacteri
408	19	95.0	212	16	Q914X9	Q914x9 pseudomonas	481	19	95.0	230	16	Q8DOF9	Q8dof9 yersinia pe
409	19	95.0	213	10	Q8H3F6	Q8h3f6 oryza sativ	482	19	95.0	230	16	Q7WQY2	Q7wqy2 bordetella
410	19	95.0	213	16	Q9KPF59	Q9kpf59 vibrio chol	483	19	95.0	230	16	Q7W204	Q7w204 bordetella
411	19	95.0	213	16	Q8YKX0	Q8ykb0 anabaena sp	484	19	95.0	230	16	Q7VUL1	Q7vul1 bordetella
412	19	95.0	213	16	Q8NTZ8	Q8ntz8 corynebacte	485	19	95.0	231	5	Q8MXA3	Q8mxa3 trypanosoma
413	19	95.0	213	16	Q87VGB	Q87vg8 pseudomonas	486	19	95.0	231	5	O81684	O81684 drosophila
414	19	95.0	213	17	Q9HR09	Q9hr09 halobacteri	487	19	95.0	231	16	Q9HXR1	Q9hxr1 pseudomonas
415	19	95.0	214	2	O54380	O54380 lactococcus	488	19	95.0	231	16	Q882T2	Q882t2 rhizobium l
416	19	95.0	214	2	Q8KP00	Q8kp00 gamma-prote	489	19	95.0	231	16	Q9ZE94	Q9ze94 listeria in
417	19	95.0	215	2	Q9RC94	Q9rc84 bacillus ha	490	19	95.0	231	16	Q8YXG7	Q8yxg7 brucella me
418	19	95.0	215	2	Q9Z480	Q9z480 agrobacteri	491	19	95.0	231	16	Q8Y9H1	Q8y9h1 listeria mo
419	19	95.0	215	10	O22136	O22136 arabidopsis	492	19	95.0	231	16	Q8XUT7	Q8xut7 ralstonia s
420	19	95.0	215	16	Q98PH0	Q98ph0 rhizobium l	493	19	95.0	231	16	Q8PQL3	Q8pql3 corynebacte
421	19	95.0	215	16	Q8PQ35	Q8pq35 xanthomonas	494	19	95.0	232	2	Q9RBD1	Q9rbd1 arthrobacte
422	19	95.0	216	2	Q9REB2	Q9reb2 arthrobacte	495	19	95.0	232	16	Q9A6A4	Q9a6a4 caulobacter
423	19	95.0	216	14	P97979	P97979 unidentified	496	19	95.0	232	16	P96241	P96241 mycobacteri
424	19	95.0	216	16	Q92Q81	Q92q81 rhizobium m	497	19	95.0	232	16	Q7TVJ7	Q7tvj7 mycobacteri
425	19	95.0	216	16	Q88MA0	Q88ma0 pseudomonas	498	19	95.0	232	17	Q8PVL0	Q8pvl0 methanosarc
426	19	95.0	217	16	Q88GA2	Q88ga2 pseudomonas	499	19	95.0	234	2	Q9RBC9	Q9rbcb9 arthrobacte
427	19	95.0	217	17	Q8UIA8	Q8uia8 pyrococcus	500	19	95.0	234	5	Q9N327	Q9n327 caenorhabdi
428	19	95.0	218	2	Q9REX4	Q9rex4 arthrobacte	501	19	95.0	234	10	Q9FQC4	Q9fqc4 zea mays (m
429	19	95.0	218	10	Q40E30	Q40e30 oryza sativ	502	19	95.0	234	10	Q94K12	Q94k12 arabidopsis
430	19	95.0	218	10	Q9L2I5	Q9l2y5 arabidopsis	503	19	95.0	234	10	Q8XIF8	Q8xif8 arabidopsis
431	19	95.0	218	16	Q91470	Q91470 pseudomonas	504	19	95.0	234	16	Q8YSQ0	Q8ysq0 anabaena sp
432	19	95.0	218	16	Q8Z5S1	Q8zes1 yersinia pe	505	19	95.0	234	16	Q8XBP4	Q8xbp4 escherichia
433	19	95.0	218	16	Q88NB6	Q88nb6 pseudomonas	506	19	95.0	234	16	Q81DL3	Q81dl3 bacillus ce
434	19	95.0	219	16	Q92B71	Q92e71 listeria in	507	19	95.0	235	2	Q9AKL9	Q9akl9 rickettsia
435	19	95.0	219	16	Q69984	Q69984 streptomyce	508	19	95.0	235	2	Q9AKB8	Q9akb8 rickettsia
436	19	95.0	219	16	Q8ZKE5	Q8zkes streptomyce	509	19	95.0	235	2	Q9AKG3	Q9akg3 rickettsia
437	19	95.0	219	16	Q7U6G1	Q7ugg1 rhodospirill	510	19	95.0	235	2	Q849C9	Q849c9 streptomyce
438	19	95.0	220	17	Q26842	Q26842 methanobact	511	19	95.0	235	16	Q8W6X8	Q8w6x8 cyanophaga
439	19	95.0	220	17	Q9HQA1	Q9hqa1 halobacteri	512	19	95.0	235	16	Q92XB1	Q92xb1 rhizobium m
440	19	95.0	221	4	Q96GF5	Q96gf5 homo sapien	513	19	95.0	235	16	Q8Z1J4	Q8z1j4 rickettsia
441	19	95.0	221	16	Q9KAB1	Q9kab1 bacillus ha	514	19	95.0	235	16	Q8XYZ8	Q8xyz8 ralstonia s
442	19	95.0	221	16	Q93J34	Q93j34 streptomyce	515	19	95.0	236	2	Q9UPA0	Q9upa0 rhodocyclu
443	19	95.0	221	16	Q82CX2	Q82cx2 streptomyce	516	19	95.0	236	10	Q9LVA8	Q9lva8 arabidopsis
444	19	95.0	222	2	Q9RBD9	Q9rbd9 arthrobacte	517	19	95.0	236	10	Q8SK05	Q8sk05 arabidopsis
445	19	95.0	222	16	Q8F9P8	Q8f9p8 leptospira	518	19	95.0	236	16	Q8UBG5	Q8ubg5 agrobacteri
446	19	95.0	223	5	O8SUL0	O8sul0 encephalito	519	19	95.0	237	2	Q9RBC0	Q9rbco arthrobacte
447	19	95.0	223	10	Q94B57	Q94b57 arabidopsis	520	19	95.0	237	2	Q84B84	Q84b84 vibrio fisc
448	19	95.0	223	16	Q9ZBP8	Q9zbp8 streptomyce	521	19	95.0	237	16	Q9KFS9	Q9kfs9 bacillus ha
449	19	95.0	223	16	Q883K9	Q883k9 pseudomonas	522	19	95.0	237	16	Q98N57	Q98n57 rhizobium l
450	19	95.0	223	16	Q8XGZ1	Q8xgz1 salmonella	523	19	95.0	237	16	Q97IV4	Q97iv4 clostridium
451	19	95.0	224	3	Q03362	Q03362 saccharomyc	524	19	95.0	237	16	Q8ELK7	Q8elk7 oceanobacil
452	19	95.0	224	4	Q92799	Q92799 homo sapien	525	19	95.0	237	17	Q8TMA2	Q8tma2 methanosarc
453	19	95.0	224	11	Q99NG0	Q99ng0 rattus norv	526	19	95.0	238	2	Q9RBP4	Q9rbp4 rhodococcus
454	19	95.0	224	16	Q8Y8J7	Q8y8j7 listeria mo	527	19	95.0	238	16	Q9CIV5	Q9civ5 lactococcus

528	19	95.0	238	16	Q9L246	Q9L246 streptomyce	601	19	95.0	253	16	Q8X4N5	Q8X4N5 escherichia
529	19	95.0	239	3	Q9UVW7	Q9UVW7 candida alb	602	19	95.0	253	16	Q9AK07	Q9AK07 streptomyce
530	19	95.0	239	16	Q836K1	Q836K1 enterococcu	603	19	95.0	253	16	Q8FE98	Q8FE98 escherichia
531	19	95.0	239	16	Q82UL5	Q82UL5 nitrosomona	604	19	95.0	253	16	Q83QB3	Q83QB3 shigella fl
532	19	95.0	239	17	Q82I13	Q82I13 pyrococcus	605	19	95.0	253	17	Q9YB95	Q9YB95 aeropyrum p
533	19	95.0	240	17	Q30144	Q30144 archaeoglob	606	19	95.0	254	2	Q9REK0	Q9REK0 athrobacte
534	19	95.0	241	2	Q7WTF7	Q7WTF7 streptomyce	607	19	95.0	254	2	Q82501	Q82501 clostridium
535	19	95.0	241	10	Q9SVW3	Q9SVW3 arabidopsis	608	19	95.0	254	8	Q8RW7	Q8RW7 guillardia
536	19	95.0	241	10	Q7XR82	Q7XR82 oryza sativ	609	19	95.0	254	10	Q94A55	Q94A55 arabidopsis
537	19	95.0	241	16	Q7WJN5	Q7WJN5 bordetella	610	19	95.0	254	10	Q22536	Q22536 oryza sativ
538	19	95.0	241	16	Q7WAI8	Q7WAI8 bordetella	611	19	95.0	254	10	Q9ATW0	Q9ATW0 zea mays (m
539	19	95.0	242	5	Q81683	Q81683 drosophila	612	19	95.0	254	16	Q8ZGH5	Q8ZGH5 yersinia pe
540	19	95.0	242	5	Q81682	Q81682 drosophila	613	19	95.0	255	2	Q9RBE0	Q9RBE0 athrobacte
541	19	95.0	242	16	Q8Y0T0	Q8Y0T0 ralstonia s	614	19	95.0	255	8	Q950V2	Q950V2 tetrachyana
542	19	95.0	242	16	Q8DH96	Q8DH96 synchococc	615	19	95.0	255	16	Q89JK6	Q89JK6 bradyrhizob
543	19	95.0	242	16	Q8CYH8	Q8CYH8 streptococc	616	19	95.0	256	2	Q9RBD5	Q9RBD5 athrobacte
544	19	95.0	243	2	Q9AN57	Q9AN57 bradyrhizob	617	19	95.0	256	8	Q9T7M4	Q9T7M4 tetrachyana
545	19	95.0	243	16	Q8G0Y5	Q8G0Y5 brucella su	618	19	95.0	256	10	Q94HB4	Q94HB4 oryza sativ
546	19	95.0	244	6	Q8HXL4	Q8HXL4 sus scrofa	619	19	95.0	256	10	Q65325	Q65325 arabidopsis
547	19	95.0	244	16	Q8DSN5	Q8DSN5 streptococc	620	19	95.0	256	10	Q7XFI7	Q7XFI7 oryza sativ
548	19	95.0	244	16	Q7UQX9	Q7UQX9 rhodopirell	621	19	95.0	256	10	Q9WN21	Q9WN21 frangipani
549	19	95.0	245	10	Q940J5	Q940J5 arabidopsis	622	19	95.0	257	10	Q8GUK5	Q8GUK5 arabidopsis
550	19	95.0	245	10	Q7XK61	Q7XK61 oryza sativ	623	19	95.0	257	11	Q9CP85	Q9CP85 mus musculu
551	19	95.0	246	6	Q28535	Q28535 mustela vis	624	19	95.0	257	11	Q99JB5	Q99JB5 rhizobium m
552	19	95.0	246	16	Q8P151	Q8P151 xanthomonas	625	19	95.0	257	16	Q92L02	Q92L02 corynebacte
553	19	95.0	246	16	Q8P6V1	Q8P6V1 xanthomonas	626	19	95.0	257	16	Q8NL77	Q8NL77 streptomyce
554	19	95.0	246	16	Q8RDB0	Q8RDB0 streptomyce	627	19	95.0	257	16	Q9X7Q4	Q9X7Q4 athrobacte
555	19	95.0	246	16	Q8FM47	Q8FM47 corynebacte	628	19	95.0	258	2	Q9RBB1	Q9RBB1 streptomyce
556	19	95.0	246	16	Q89MY3	Q89MY3 bradyrhizob	629	19	95.0	258	13	Q7S2S8	Q7S2S8 xenopus lae
557	19	95.0	246	16	Q82AC3	Q82AC3 streptomyce	630	19	95.0	258	16	Q9HZN9	Q9HZN9 pseudomonas
558	19	95.0	247	2	Q48830	Q48830 lactobacill	631	19	95.0	258	16	Q9X8L2	Q9X8L2 streptomyce
559	19	95.0	247	16	Q88ZC3	Q88ZC3 lactobacill	632	19	95.0	258	16	Q7WCV7	Q7WCV7 bordetella
560	19	95.0	248	5	Q97008	Q97008 leishmania	633	19	95.0	258	16	Q7VZK8	Q7VZK8 bordetella
561	19	95.0	248	10	Q949B9	Q949B9 oryza sativ	634	19	95.0	258	16	Q7VZK8	Q7VZK8 bordetella
562	19	95.0	248	10	Q99935	Q99935 helianthus	635	19	95.0	258	17	Q9HLA1	Q9HLA1 thermoplasm
563	19	95.0	248	10	Q81216	Q81216 zea mays (m	636	19	95.0	259	2	Q9JSP2	Q9JSP2 eubacterium
564	19	95.0	248	10	Q84RL6	Q84RL6 zea mays (m	637	19	95.0	259	2	Q8VMN6	Q8VMN6 pseudomonas
565	19	95.0	248	10	Q7XAL6	Q7XAL6 oryza sativ	638	19	95.0	259	13	Q7ZTI9	Q7ZTI9 xenopus lae
566	19	95.0	248	16	Q8ZG59	Q8ZG59 yersinia pe	639	19	95.0	259	16	P74082	P74082 synchocyst
567	19	95.0	248	16	Q8A971	Q8A971 bacteroides	640	19	95.0	259	16	P74082	P74082 campylobact
568	19	95.0	248	17	Q9YCF6	Q9YCF6 aeropyrum p	641	19	95.0	259	16	Q9PJ15	Q9PJ15 campylobact
569	19	95.0	249	5	Q62042	Q62042 caenorhabdi	642	19	95.0	259	16	Q98K61	Q98K61 rhizobium l
570	19	95.0	249	10	Q9LXK3	Q9LXK3 hordeum vul	643	19	95.0	259	16	Q7U9R3	Q7U9R3 synchococc
571	19	95.0	249	10	Q9ATL9	Q9ATL9 zea mays (m	644	19	95.0	259	17	Q8ZT71	Q8ZT71 pyrobaculum
572	19	95.0	249	10	Q7XU33	Q7XU33 oryza sativ	645	19	95.0	260	2	Q9RBA7	Q9RBA7 athrobacte
573	19	95.0	249	12	Q919L4	Q919L4 culex nigri	646	19	95.0	260	2	Q8G976	Q8G976 pseudomonas
574	19	95.0	249	16	Q8PKD5	Q8PKD5 chlamydia m	647	19	95.0	260	2	Q849Q8	Q849Q8 pseudomonas
575	19	95.0	249	16	Q8PG49	Q8PG49 xanthomonas	648	19	95.0	260	16	Q8YVJ0	Q8YVJ0 arabidopsis
576	19	95.0	249	16	Q8FIP0	Q8FIP0 leptospira	649	19	95.0	260	16	Q8F545	Q8F545 corynebacte
577	19	95.0	249	16	Q93N41	Q93N41 coxiella bu	650	19	95.0	260	16	Q81EA0	Q81EA0 bacillus ce
578	19	95.0	249	16	Q82M75	Q82M75 streptomyce	651	19	95.0	260	16	Q7WNS7	Q7WNS7 bordetella
579	19	95.0	249	17	Q29512	Q29512 archaeoglob	652	19	95.0	260	16	Q7WU07	Q7WU07 bordetella
580	19	95.0	250	10	Q9ATL8	Q9ATL8 zea mays (m	653	19	95.0	260	16	Q7WV38	Q7WV38 bordetella
581	19	95.0	250	11	Q84JF6	Q84JF6 arabidopsis	654	19	95.0	261	2	Q66247	Q66247 actinobacil
582	19	95.0	250	11	Q8C4H6	Q8C4H6 mus musculu	655	19	95.0	261	2	Q9XDR0	Q9XDR0 actinobacil
583	19	95.0	250	16	Q25228	Q25228 helicobacte	656	19	95.0	261	3	Q9P8F1	Q9P8F1 glomus moss
584	19	95.0	251	2	Q93A69	Q93A69 gamma-prote	657	19	95.0	261	3	Q9P8E9	Q9P8E9 glomus moss
585	19	95.0	251	5	Q8SVK8	Q8SVK8 encephalito	658	19	95.0	261	6	Q951N6	Q951N6 macaca fasc
586	19	95.0	251	11	Q9D424	Q9D424 mus musculu	659	19	95.0	261	11	Q8K2L3	Q8K2L3 mus musculu
587	19	95.0	251	17	Q23320	Q23320 archaeoglob	660	19	95.0	261	11	Q9DCE1	Q9DCE1 mus musculu
588	19	95.0	252	2	Q03540	Q03540 escherichia	661	19	95.0	261	11	Q9JMJ6	Q9JMJ6 mus musculu
589	19	95.0	252	9	Q8HA20	Q8HA20 bacterioph	662	19	95.0	261	11	Q9BH40	Q9BH40 mus musculu
590	19	95.0	252	10	Q94CS9	Q94CS9 oryza sativ	663	19	95.0	261	11	Q98W5	Q98W5 bacillus ha
591	19	95.0	252	10	Q7XA62	Q7XA62 oryza sativ	664	19	95.0	261	16	Q98X77	Q98X77 rhizobium l
592	19	95.0	252	16	Q9A534	Q9A534 caulobacter	665	19	95.0	261	16	Q8UCAL	Q8UCAL agrobacteri
593	19	95.0	252	16	Q8A066	Q8A066 lactobacill	666	19	95.0	261	16	Q8U893	Q8U893 agrobacteri
594	19	95.0	252	16	Q7UT61	Q7UT61 rhodopirell	667	19	95.0	261	16	Q87YH4	Q87YH4 pseudomonas
595	19	95.0	252	17	Q97VW6	Q97VW6 sulfolobus	668	19	95.0	261	16	Q7WPP1	Q7WPP1 bordetella
596	19	95.0	252	17	Q97VW6	Q97VW6 sulfolobus	669	19	95.0	261	16	Q7WBP9	Q7WBP9 gamma-prote
597	19	95.0	253	2	Q8KH15	Q8KH15 pectobacter	670	19	95.0	262	3	Q9P8B8	Q9P8B8 glomus moss
598	19	95.0	253	5	Q95U55	Q95U55 drosophila	671	19	95.0	262	3	Q9WZP2	Q9WZP2 thermotoga
599	19	95.0	253	10	Q8W371	Q8W371 oryza sativ	672	19	95.0	262	16	Q8EDY1	Q8EDY1 shewanella
600	19	95.0	253	10	Q7XD92	Q7XD92 oryza sativ	673	19	95.0	262	17	Q8U0A9	Q8U0A9 pyrococcus

674	19	95.0	263	16	Q8D064	Q8d064 yersinia pe	747	19	95.0	276	16	Q8Z016	Q8z016 anabaena sp
675	19	95.0	263	16	Q8A0Q5	Q8a0q5 bacteroides	748	19	95.0	276	16	Q8Pp23	Q8pp23 xanthomonas
676	19	95.0	263	16	Q7Wn22	Q7wn22 bordetella	749	19	95.0	276	16	Q881P0	Q881p0 pseudomonas
677	19	95.0	263	16	Q7WbJ9	Q7wbj9 bordetella	750	19	95.0	277	2	Q85287	Q85287 sphingomona
678	19	95.0	264	4	Q8TDN7	Q8tdn7 homo sapien	751	19	95.0	277	2	P72544	P72544 synechococc
679	19	95.0	264	10	Q940D9	Q940d9 pinus taeda	752	19	95.0	277	16	Q8P8X6	Q8px6 xanthomonas
680	19	95.0	265	4	Q8TDC1	Q8tdc1 homo sapien	753	19	95.0	278	16	Q9JXZ5	Q9jxz5 neisseria m
681	19	95.0	265	5	O15953	O15953 drosophila	754	19	95.0	278	16	Q9JVK5	Q9jvk5 neisseria m
682	19	95.0	265	10	Q8G3T6	Q8g3t6 cryza sativ	755	19	95.0	278	16	Q83JY4	Q83jy4 treponema p
683	19	95.0	265	11	Q91ZX8	Q91zx8 rattus norv	756	19	95.0	278	16	Q8E5D6	Q8ed6 streptococc
684	19	95.0	265	17	Q978W2	Q978w2 thermoplasm	757	19	95.0	279	16	Q8E5D6	Q8ed6 streptococc
685	19	95.0	265	17	Q978W2	Q978w2 thermoplasm	758	19	95.0	279	16	Q8E5D6	Q8ed6 streptococc
686	19	95.0	267	2	O05364	O05364 actinobacil	759	19	95.0	279	16	Q8E5D6	Q8ed6 streptococc
687	19	95.0	267	2	Q83Y72	Q83y72 gamma-prote	760	19	95.0	279	17	Q7VTF9	Q7vtf9 bordetella
688	19	95.0	267	10	Q8GZD2	Q8gzd2 zea mays (m	761	19	95.0	280	2	P71101	P71101 curtopbacter
689	19	95.0	267	11	Q8BWM5	Q8bw5 mus musculu	762	19	95.0	280	2	Q8GFG7	Q8gfg7 rhodococcus
690	19	95.0	267	16	Q8A8V3	Q8a8v3 caulobacter	763	19	95.0	280	10	Q9LGD0	Q9lgd0 cryza sativ
691	19	95.0	267	16	Q8U757	Q8u757 agrobacteri	764	19	95.0	280	16	Q92LFF	Q92lff rhizobium m
692	19	95.0	267	17	Q88HZ2	Q88hz2 pseudomonas	765	19	95.0	280	16	Q82LFF	Q82lff rhizobium m
693	19	95.0	267	17	Q8ZXK1	Q8zxk1 pyrobaculum	766	19	95.0	280	16	Q8NRY7	Q8nry7 corynebacte
694	19	95.0	268	5	Q8T4T4	Q8t4t4 aedes aegyp	767	19	95.0	280	17	Q8TU99	Q8tu99 methanosarc
695	19	95.0	268	5	Q8NOR8	Q8nor8 aedes aegyp	768	19	95.0	281	2	Q9ADV7	Q9adv7 ehrlichia c
696	19	95.0	268	5	Q8RIG0	Q8rig0 aedes aegyp	769	19	95.0	281	10	Q9ZD17	Q9zd17 arabisidopsis
697	19	95.0	268	5	Q8T4T5	Q8t4t5 aedes aegyp	770	19	95.0	281	10	O04035	O04035 arabisidopsis
698	19	95.0	268	5	Q8T4T5	Q8t4t5 aedes aegyp	771	19	95.0	281	16	Q82M07	Q82m07 streptomyc
699	19	95.0	268	5	Q16900	Q16900 aedes aegyp	772	19	95.0	281	16	Q82M07	Q82m07 streptomyc
700	19	95.0	268	10	Q8W189	Q8w189 cryza sativ	773	19	95.0	281	16	Q7UC85	Q7uc85 sulfobolus
701	19	95.0	268	16	Q8D0N4	Q8d0n4 yersinia pe	774	19	95.0	281	17	Q962Y6	Q962y6 sulfobolus
702	19	95.0	269	4	Q8TBI5	Q8tb15 homo sapien	775	19	95.0	282	2	Q8KH99	Q8kha9 pseudomonas
703	19	95.0	269	11	Q91VY8	Q91vy8 mus musculu	776	19	95.0	282	2	O68913	O68913 streptomyc
704	19	95.0	269	11	Q7TN36	Q7tn36 rattus norv	777	19	95.0	282	16	Q8KGP1	Q8kgp1 bacillus ha
705	19	95.0	269	16	Q8FS10	Q8fs10 corynebacte	778	19	95.0	282	16	Q8KGP1	Q8kgp1 bacillus ha
706	19	95.0	269	16	Q7WFF26	Q7wff26 bordetella	779	19	95.0	282	16	Q92688	Q92688 clostridium
707	19	95.0	269	16	Q7W3P9	Q7w3p9 bordetella	780	19	95.0	282	16	Q8DZN8	Q8dzn8 streptococc
708	19	95.0	269	16	Q7VST6	Q7vst6 bordetella	781	19	95.0	282	16	Q89176	Q89176 bradyrhizob
709	19	95.0	269	17	Q57836	Q57836 pyrococcus	782	19	95.0	282	16	Q82DZ5	Q82dz5 streptomyc
710	19	95.0	270	16	Q8NTB3	Q8ntb3 corynebacte	783	19	95.0	283	2	O85983	O85983 sphingomona
711	19	95.0	270	16	Q86443	Q86443 pseudomonas	784	19	95.0	283	2	Q83VK3	Q83vk3 sphingomona
712	19	95.0	271	6	Q9N2J4	Q9n2j4 canis famil	785	19	95.0	284	3	O13848	O13848 schizosacch
713	19	95.0	271	8	O21312	O21312 tetrahymena	786	19	95.0	284	3	O13848	O13848 schizosacch
714	19	95.0	271	16	Q7UNX6	Q7unx6 rhodopirell	787	19	95.0	284	4	Q9HAB8	Q9hab8 homo sapien
715	19	95.0	271	16	Q7UMF0	Q7umf0 rhodopirell	788	19	95.0	284	16	Q8DTC4	Q8dte4 streptococc
716	19	95.0	272	2	Q8GCV5	Q8gcv5 lactobacill	789	19	95.0	284	16	Q8A5T6	Q8ast6 bacteroides
717	19	95.0	272	5	Q8GFH0	Q8gfh0 rhodococcus	790	19	95.0	285	10	Q9SHW9	Q9shw9 arabisidopsis
718	19	95.0	272	5	Q17131	Q17131 babesia equ	791	19	95.0	285	12	Q9PVX4	Q9pvx4 xestia c-ni
719	19	95.0	272	5	Q9TYB0	Q9tyb0 babesia equ	792	19	95.0	285	15	O87X11	O87x11 pseudomonas
720	19	95.0	272	5	Q9TYB1	Q9tyb1 babesia equ	793	19	95.0	286	2	O05327	O05327 sphingomona
721	19	95.0	272	5	Q9TYA7	Q9tya7 babesia equ	794	19	95.0	286	13	Q7ZVV6	Q7zvv6 brachydania
722	19	95.0	272	5	Q9TYB2	Q9tyb2 babesia equ	795	19	95.0	286	16	Q8ZPM8	Q8zpm8 salmonella
723	19	95.0	272	5	Q9TYA9	Q9tya9 babesia equ	796	19	95.0	286	16	Q8Z521	Q8ze21 yersinia pe
724	19	95.0	272	5	Q9TYA8	Q9tya8 babesia equ	797	19	95.0	286	16	Q8DWF0	Q8dwf0 streptococc
725	19	95.0	272	5	O46365	O46365 babesia equ	798	19	95.0	286	16	Q883F1	Q883f1 pseudomonas
726	19	95.0	272	5	Q9TYB3	Q9tyb3 babesia equ	799	19	95.0	287	5	Q81410	Q81410 caenorhabdi
727	19	95.0	272	5	O45145	O45145 caenorhabdi	800	19	95.0	287	16	O8X649	O8x649 escherichia
728	19	95.0	273	2	Q9F5J1	Q9f5j1 streptomyc	801	19	95.0	287	16	O86719	O86719 streptomyc
729	19	95.0	273	5	Q20243	Q20243 caenorhabdi	802	19	95.0	287	16	O86719	O86719 streptomyc
730	19	95.0	273	11	Q8R4X1	Q8r4x1 mus musculu	803	19	95.0	287	16	Q8FHB9	Q8fhb9 escherichia
731	19	95.0	273	16	Q9PCH8	Q9pch8 streptomyc	804	19	95.0	287	16	Q8FHB9	Q8fhb9 escherichia
732	19	95.0	273	17	Q8TVU2	Q8tvu2 pyrobaculum	805	19	95.0	287	16	Q88M34	Q88m34 pseudomonas
733	19	95.0	274	5	Q20303	Q20303 caenorhabdi	806	19	95.0	287	16	Q88M34	Q88m34 pseudomonas
734	19	95.0	274	11	Q8CDS2	Q8cds2 mus musculu	807	19	95.0	288	10	Q83KY1	Q83ky1 shigella fi
735	19	95.0	274	16	Q9XFK4	Q9xfk4 bacillus ha	808	19	95.0	288	16	O83KY1	O83ky1 shigella fi
736	19	95.0	274	16	Q9A112	Q9a112 streptococc	809	19	95.0	288	16	O8H0U7	O8hou7 arabisidopsis
737	19	95.0	274	16	Q97DB2	Q97db2 clostridium	810	19	95.0	288	16	O8H0U7	O8hou7 arabisidopsis
738	19	95.0	274	16	Q8P2P0	Q8p2p0 streptococc	811	19	95.0	288	16	O8YGG8	O8ygg8 brucella me
739	19	95.0	274	16	Q8FTL7	Q8ftl7 corynebacte	812	19	95.0	288	16	O8YGG8	O8ygg8 brucella me
740	19	95.0	274	16	Q8F5I3	Q8f5i3 leptospira	813	19	95.0	288	16	O81WQ9	O81wg9 bacillus an
741	19	95.0	274	16	Q879M6	Q879m6 streptococc	814	19	95.0	288	16	O81A22	O81a22 bacillus ce
742	19	95.0	274	16	Q7VST1	Q7vst1 bordetella	815	19	95.0	289	10	Q9ST66	Q9st66 solanum tub
743	19	95.0	275	16	Q98H12	Q98h12 rhizobium 1	816	19	95.0	289	11	Q8BMV2	Q8bm2 mus musculu
744	19	95.0	275	16	Q97F28	Q97f28 clostridium	817	19	95.0	289	16	Q8BMV2	Q8bm2 mus musculu
745	19	95.0	275	16	Q82W11	Q82w11 nitrosomona	818	19	95.0	289	16	Q8BMV2	Q8bm2 mus musculu
746	19	95.0	276	5	Q18536	Q18536 caenorhabdi	819	19	95.0	290	4	Q84HE4	Q84he4 vibrio para



820	19	95.0	291	16	Q97LH6	Q97Lh6 clostridium	893	19	95.0	306	13	Q7ZX07	Q7zx07 xenopus lae
821	19	95.0	291	16	Q84533	Q84633 chlamydia t	894	19	95.0	306	16	Q8U746	Q8u746 agrobacteri
822	19	95.0	292	5	Q33766	Q33766 sphingomonas	895	19	95.0	306	16	Q8PGN2	Q8pgn2 xanthomonas
823	19	95.0	292	5	Q16616	Q16616 caenorhabdi	896	19	95.0	306	17	Q8Z2C4	Q8z2c4 pyrobaculum
824	19	95.0	292	5	Q7Y2C5	Q7Y2c5 bacterioph	897	19	95.0	307	5	Q18327	Q18327 caenorhabdi
825	19	95.0	293	5	Q9GNV9	Q9Gnv9 leishmania	898	19	95.0	307	5	Q8STV9	Q8stv9 encephalito
826	19	95.0	293	16	Q83315	Q83315 treponema p	899	19	95.0	307	10	Q9LNR7	Q9lnr7 arabidopsis
827	19	95.0	293	16	Q88LZ3	Q88Lz3 pseudomonas	900	19	95.0	307	11	Q8VGH1	Q8vgh1 mus musculu
828	19	95.0	293	16	Q828J6	Q828j6 streptomyce	901	19	95.0	307	16	Q82PD4	Q82pd4 streptomyce
829	19	95.0	294	2	Q9RGK4	Q9RGk4 bacteroides	902	19	95.0	308	11	Q8VEG8	Q8veg8 mus musculu
830	19	95.0	294	4	Q9NRW3	Q9NRw3 homo sapien	903	19	95.0	308	16	Q9CH02	Q9ch02 lactococcus
831	19	95.0	294	10	Q8G1P1	Q8G1p1 oryza sativ	904	19	95.0	308	16	Q8NLL4	Q8nll4 corynebacte
832	19	95.0	295	2	Q9DXK6	Q9DXk6 bacteroides	905	19	95.0	308	16	Q7V947	Q7v947 prochloroco
833	19	95.0	295	9	Q854V0	Q854v0 mycobacteri	906	19	95.0	309	10	Q8T000	Q8t000 arabidopsis
834	19	95.0	295	10	Q94DE7	Q94de7 oryza sativ	907	19	95.0	309	11	Q8VGU8	Q8vgu8 mus musculu
835	19	95.0	295	16	Q9EPG2	Q9EPg2 xylella fas	908	19	95.0	309	11	Q7TRC3	Q7trc3 mus musculu
836	19	95.0	295	16	Q9KV64	Q9KV64 vibrio chol	909	19	95.0	309	16	Q9ARE1	Q9are1 mycoplasma
837	19	95.0	295	163	Q8K9N8	Q8K9n8 streptococ	910	19	95.0	309	16	Q8PYV5	Q8pyv5 methanosarc
838	19	95.0	295	163	Q8AK1	Q8Aak1 bacteroides	911	19	95.0	310	4	Q9T6E6	Q9t6e6 homo sapien
839	19	95.0	295	16	Q8AET9	Q8Aet9 xylella fas	912	19	95.0	310	10	Q9SPD6	Q9spd6 arabidopsis
840	19	95.0	295	2	Q9LC87	Q9LC87 nocardioide	913	19	95.0	310	10	Q84Z99	Q84z99 oryza sativ
841	19	95.0	295	17	Q8TU61	Q8Tu61 methanosarc	914	19	95.0	310	11	Q8K2C6	Q8k2c6 mus musculu
842	19	95.0	296	17	Q8PWS7	Q8Pws7 methanosarc	915	19	95.0	310	12	Q9EMQ9	Q9emq9 anasacta moo
843	19	95.0	297	2	Q93LQ8	Q93Lq8 klebsiella	916	19	95.0	310	12	Q9CK67	Q9ck67 pasteurella
844	19	95.0	297	10	Q82163	Q82163 arabidopsis	917	19	95.0	310	16	Q88E77	Q88e77 pseudomonas
845	19	95.0	297	16	Q8YV53	Q8Yv53 anabaena sp	918	19	95.0	310	16	Q87Y83	Q87y83 pseudomonas
846	19	95.0	297	16	Q8XG56	Q8Xg56 anabaena sp	919	19	95.0	311	11	Q8VF62	Q8vf62 mus musculu
847	19	95.0	297	16	Q8XFB0	Q8Xfb0 anabaena sp	920	19	95.0	311	10	Q9FTR1	Q9ftr1 oryza sativ
848	19	95.0	297	16	Q8R977	Q8R977 thermocanaer	921	19	95.0	311	16	Q98N28	Q98n28 rhizobium l
849	19	95.0	298	13	Q7ZV21	Q7Zv21 brachydanit	922	19	95.0	311	16	Q8NU00	Q8nu00 corynebacte
850	19	95.0	298	16	Q8U7A1	Q8U7a1 agrobacteri	923	19	95.0	311	16	Q8FU71	Q8fu71 corynebacte
851	19	95.0	298	16	Q8ELP1	Q8Elp1 oceanobacil	924	19	95.0	311	16	Q88WP5	Q88wp5 lactobacilli
852	19	95.0	298	16	Q981M9	Q981m9 clostridium	925	19	95.0	311	16	Q8HMX6	Q8hmx6 halobacteri
853	19	95.0	299	2	Q9ADV6	Q9Adv6 ehrlichia c	926	19	95.0	311	17	Q97ZS5	Q97zs5 sulfolobus
854	19	95.0	299	2	Q93H57	Q93h57 bradyrhizob	927	19	95.0	311	17	Q8ZTIS	Q8ztis pyrobaculum
855	19	95.0	299	2	Q8GF66	Q8GF66 rhodococcus	928	19	95.0	311	17	Q8ZTIS	Q8ztf9 rhodococcus
856	19	95.0	299	4	Q9NXA8	Q9NXa8 homo sapien	929	19	95.0	312	3	Q8GFG9	Q8gfg9 rhodococcus
857	19	95.0	299	4	Q9NV14	Q9NV14 homo sapien	930	19	95.0	312	3	Q12125	Q12125 saccharomyc
858	19	95.0	299	4	Q9NQ56	Q9NQ56 homo sapien	931	19	95.0	312	10	Q7XYT9	Q7xyt9 oryza sativ
859	19	95.0	299	4	Q8TC61	Q8TC61 homo sapien	932	19	95.0	312	16	Q8NQR2	Q8nqr2 corynebacte
860	19	95.0	299	4	Q9NQ48	Q9NQ48 homo sapien	933	19	95.0	313	5	Q25791	Q25791 plasmodium
861	19	95.0	299	11	Q9JHQ5	Q9Jhq5 mus musculu	934	19	95.0	313	11	Q8VGE0	Q8vge0 mus musculu
862	19	95.0	299	11	Q8CDG8	Q8cdg8 mus musculu	935	19	95.0	313	16	Q92XG9	Q92xg9 rhizobium m
863	19	95.0	299	16	Q8BRX8	Q8brx8 mus musculu	936	19	95.0	313	16	Q92AX6	Q92ax6 listeria in
864	19	95.0	299	16	Q9KKW8	Q9KKw8 vibrio chol	937	19	95.0	313	16	Q8Y6K0	Q8y6k0 listeria mo
865	19	95.0	299	16	Q8EIG6	Q8Eig6 shewanella	938	19	95.0	314	5	Q8WQX2	Q8wqx2 synechococc
866	19	95.0	299	16	Q8EIG6	Q8Eig6 shewanella	939	19	95.0	314	5	Q8WQX2	Q8wfd7 mus musculu
867	19	95.0	299	16	Q8D5V7	Q8D5v7 vibrio para	940	19	95.0	314	11	Q8VFD7	Q8vfd7 mus musculu
868	19	95.0	299	16	Q87G20	Q87g20 vibrio para	941	19	95.0	314	11	Q7TRY2	Q7try2 mus musculu
869	19	95.0	299	17	Q8ZWF0	Q8Zwf0 pyrobaculum	942	19	95.0	314	11	Q7TRY1	Q7try1 mus musculu
870	19	95.0	299	17	Q8ZTV2	Q8Ztv2 pyrobaculum	943	19	95.0	314	11	Q7TRY4	Q7try4 mus musculu
871	19	95.0	300	2	Q8GF68	Q8GF68 rhodococcus	944	19	95.0	314	16	Q8YKH7	Q8ykh7 anabaena sp
872	19	95.0	300	16	Q8EJ78	Q8Ejt8 streptococ	945	19	95.0	314	16	Q8FL19	Q8fl19 xanthomonas
873	19	95.0	300	16	Q8DY72	Q8DY72 streptococ	946	19	95.0	314	16	Q8P9B2	Q8p9b2 xanthomonas
874	19	95.0	300	16	Q7W5F1	Q7W5f1 bordetella	947	19	95.0	314	16	Q8FVU5	Q8fvu5 brucella su
875	19	95.0	301	5	Q3IHK7	Q3Ihk7 babesia div	948	19	95.0	314	16	Q88HX8	Q88hx8 pseudomonas
876	19	95.0	301	10	Q93X14	Q93x14 oryza sativ	949	19	95.0	314	16	Q81JUS1	Q81js1 bacillus an
877	19	95.0	301	16	Q7VY47	Q7Vv47 bordetella	950	19	95.0	314	16	Q81JG8	Q81jg8 bacillus an
878	19	95.0	301	17	Q8ZAP5	Q8Zap5 pyrobaculum	951	19	95.0	314	16	Q81JG8	Q81jg8 bacillus an
879	19	95.0	302	8	Q9SGH8	Q9SGh8 arabidopsis	952	19	95.0	314	17	Q82600	Q82ab9 homo sapien
880	19	95.0	302	10	Q9SGS4	Q9SGs4 arabidopsis	953	19	95.0	315	4	Q8TBB9	Q8tab9 pan troglod
881	19	95.0	302	10	Q8LDJ9	Q8LDj9 arabidopsis	954	19	95.0	315	6	Q9TSM8	Q9tem8 macaca fasc
882	19	95.0	302	16	Q8KAL5	Q8Kal5 chlorobium	955	19	95.0	315	6	Q9TSM8	Q9tem7 macaca fasc
883	19	95.0	303	5	Q8M275	Q8M275 drosophila	956	19	95.0	315	6	Q9TSM8	Q9tu88 gorilla gor
884	19	95.0	303	5	Q8IPV4	Q8IPv4 drosophila	957	19	95.0	315	6	Q9TUA0	Q9tua0 pan troglod
885	19	95.0	303	16	Q8NTU5	Q8Ntu5 corynebacte	958	19	95.0	315	6	Q9TUA0	Q9tu85 gorilla gor
886	19	95.0	303	16	Q7WHF7	Q7Whf7 bordetella	959	19	95.0	315	10	Q8W5L7	Q8w5l7 oryza sativ
887	19	95.0	303	16	Q7W9B0	Q7W9b0 bordetella	960	19	95.0	315	10	Q7XH71	Q7xh71 oryza sativ
888	19	95.0	303	16	Q7VVR7	Q7Vvr7 bordetella	961	19	95.0	315	11	Q9EQ93	Q9eq93 mus musculu
889	19	95.0	304	2	Q3IAC8	Q3Iac8 pseudomonas	962	19	95.0	315	5	Q8W525	Q8w525 trypanosoma
890	19	95.0	304	10	Q8LBS8	Q8Lbs8 arabidopsis	963	19	95.0	316	10	Q9SP22	Q9spe2 malus domes
891	19	95.0	304	11	Q8BN03	Q8Bn03 mus musculu	964	19	95.0	316	16	Q7UKU5	Q7uku5 rhodopirell
892	19	95.0	305	2	Q84EL1	Q84el1 campylobact	965	19	95.0	317	5	Q9U375	Q9u375 caenorhabdi

966 19 95.0 317 9 Q854X2 mycobacteri  
967 19 95.0 317 10 Q81AG7 arabidopsis  
968 19 95.0 317 11 P97827 rattus norv  
969 19 95.0 317 11 Q322J2 cricetus  
970 19 95.0 317 16 Q812Y4 bacillus ce  
971 19 95.0 317 17 Q30168 archaeglob  
972 19 95.0 318 2 Q9X5F7 zymomonas m  
973 19 95.0 318 2 Q83UD0 neisseria g  
974 19 95.0 318 2 Q7WZ93 nomomuraea  
975 19 95.0 318 3 Q9C135 mortierella  
976 19 95.0 318 10 Q8LCB9 arabidopsis  
977 19 95.0 318 10 Q7XE27 oryza sativ  
978 19 95.0 318 11 Q9CZ21 mus musculu  
979 19 95.0 318 11 Q8VGA9 mus musculu  
980 19 95.0 318 16 Q8FEC0 escherichia  
981 19 95.0 318 16 Q83LQ1 shigella fl  
982 19 95.0 318 17 Q8TT71 methanosarc  
983 19 95.0 318 17 Q8TQ14 methanosarc  
984 19 95.0 319 3 Q9C0Z8 schizosacch  
985 19 95.0 319 5 Q9STC9 drosophila  
986 19 95.0 319 10 Q9SW32 arabidopsis  
987 19 95.0 319 16 Q8FTS9 corynebacte  
988 19 95.0 319 16 Q83E34 coxiella bu  
989 19 95.0 319 17 Q9HK71 thermoplasm  
990 19 95.0 319 17 Q8TR50 methanosarc  
991 19 95.0 320 10 Q8LD47 arabidopsis  
992 19 95.0 320 10 Q8LFT4 arabidopsis  
993 19 95.0 320 10 Q8LKN3 arabidopsis  
994 19 95.0 320 10 Q9LUF3 arabidopsis  
995 19 95.0 320 10 Q9C972 arabidopsis  
996 19 95.0 320 10 Q84WV9 arabidopsis  
997 19 95.0 320 10 Q9CAH5 arabidopsis  
998 19 95.0 321 2 Q7X1G1 leptospiril  
999 19 95.0 321 4 Q8N6A0 homo sapien  
1000 19 95.0 322 2 Q7WX56 alcaligenes

## ALIGNMENTS

## RESULT 1

99TS18  
AC Q9TS18 PRELIMINARY; PRT; 20 AA.  
QT 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Cytotoxin-binding protein (Fragment).  
XS Cryptolagus cuniculus (Rabbit).  
XC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
XC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
XX NCBI\_TaxID=9986;  
[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=94039134; PubMed=763466;  
RA Lutz F., Mohr M., Grimmig M., Leidolf R., Linder D.;  
RT "Pseudomonas aeruginosa cytoxin-binding protein in rabbit  
erythrocyte membranes. An oligomer of 28 kDa with similarity to  
transmembrane channel proteins".  
RL Eur. J. Biochem. 217:1123-1128(1993).  
DR PIR; S39049; S39049.  
DR HSSP; P47865; 1U4N.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; P:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000425; MTP.  
DR Pfam; PF00230; MTP; 1  
SQ SEQUENCE 20 AA; 2414 MW; 40B82D7DB5283D2D CRC64;

Query Match 95.0%; Score 19; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 14 VAEF 17

## RESULT 2

Q9R5C0 PRELIMINARY; PRT; 22 AA.  
ID Q9R5C0  
AC Q9R5C0  
DT 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMELrel. 14, Last annotation update)  
DE Cytochrome P-460 (Fragment).  
OS Nitrosomonas europaea.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
OC Nitrosomonadaceae; Nitrosomonas.  
NCBI\_TaxID=915;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93224468; PubMed=8385668;  
RA McTavish H., LaQuier F., Arciero D., Logan M., Mundfrom G.,  
Fuchs J.A., Hooper A.B.;  
RT "Multiple copies of genes coding for electron transport proteins in  
the bacterium Nitrosomonas europaea".  
RL J. Bacteriol. 175:2445-2447(1993).  
SQ SEQUENCE 22 AA; 2549 MW; C22664F5EEICE75F CRC64;

Query Match 95.0%; Score 19; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 3 VAEF 6

## RESULT 3

P82214 PRELIMINARY; PRT; 30 AA.  
ID P82214  
AC P82214  
DT 01-OCT-2001 (TREMELrel. 18, Created)  
DT 01-OCT-2001 (TREMELrel. 18, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Myosin regulatory light chain 2 (MLC-2) (Fragment).  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
OC Bombycidae; Bombyx.  
NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE.  
RX STRAIN=XINHANG X KEMING; TISSUE=Body wall, and Fat body;  
RX MEDLINE=21177481; PubMed=11280994;  
RA Zhong B.X.;  
RT "Protein database for several tissues derived from five instar of  
silkworm".  
RL I Chuan Hsueh Pao 28:217-224(2001).  
CC -1- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS  
(BY SIMILARITY).  
CC -1- MISCELLANEOUS: THIS CHAIN BINDS CALCIUM (BY SIMILARITY).  
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.  
DR GO; GO:0016459; C:myosin; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0007517; P:muscle development; IEA.  
DR InterPro; IPR02048; EF-hand.  
DR PROSITE; PS00018; EF HAND; PARTIAL.  
KW Myosin; Calcium-binding; Muscle protein.  
FT CA BIND 25 >30 BY SIMILARITY.  
FT NON TER 30  
SQ SEQUENCE 30 AA; 3369 MW; 5B5287CE9EF538F3 CRC64;

Query Match 95.0%; Score 19; DB 5; Length 30;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|  
|  
|  
13 VAEF 16

SUITS 4

5518

O25518 PRELIMINARY; PRT; 33 AA.

O25518, 1998 (TRENBLrel. 05, Created)

01-JAN-1998 (TRENBLrel. 05, Last sequence update)

01-JUN-2003 (TRENBLrel. 24, Last annotation update)

Hypothetical protein HP0847.

HP0847.

Helicobacter pylori (Campylobacter pylori).

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

Helicobacteraceae; Helicobacter.

NCBI\_TaxID=210;

[1]

SEQUENCE FROM N.A.

STRAIN=26695 / ATCC 700392;

MEDLINE=97394467; PubMed=9252185;

Tomb J.-E., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,

Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,

Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,

Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,

McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,

Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,

Corton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,

Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,

Venter J.C.;

"The complete genome sequence of the gastric pathogen Helicobacter

pylori."

Nature 388:539-547(1997).

EMBL: AE000596; AAD07920.1; --

PIR: G64625; G64625.

TIGR: HP0847; --

Hypothetical protein; Complete proteome.

SEQUENCE 33 AA; 3704 MW; FA3F52631C0DB943 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 33;

Best Local Similarity 100.0%; Pred. No. 5.1e-02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5

|  
|  
|  
15 VAEF 18

SUITS 5

ZZ19

O82Z19 PRELIMINARY; PRT; 35 AA.

O82Z19, 1998 (TRENBLrel. 20, Created)

01-MAR-2002 (TRENBLrel. 20, Last sequence update)

01-JUN-2003 (TRENBLrel. 24, Last annotation update)

Hypothetical protein PAE0485.

PAE0485.

Pyrobaculum aerophilum.

Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;

Thermoproteaceae; Pyrobaculum.

NCBI\_TaxID=13773;

[1]

SEQUENCE FROM N.A.

STRAIN=IM2 / ATCC 51768 / DSM 7523;

MEDLINE=21664397; PubMed=11792869;

Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,

Miller J.H.;

"Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum

aerophilum."

Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).

DR EMBL; AE009770; AAL62822.1; --

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 35 AA; 3934 MW; A324B4F94F6A14BB CRC64;

Query Match 95.0%; Score 19; DB 17; Length 35;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

|  
|  
|  
18 VAEF 21

RESULT 6

Q9TQR7

ID Q9TQR7 PRELIMINARY; PRT; 36 AA.

AC Q9TQR7;

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE Transferrin (Fragment).

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OC NCBI\_TaxID=9796;

OC [1]

SEQUENCE FROM N.A.

RP Giffard J.M., Brandon R.B., Bell T.K.;

RA "Further identification of single nucleotide polymorphisms in the

RT equine transferrin gene."

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF185729; AAF05508.1; --

DR EMBL; AF185722; AAF05501.1; --

DR EMBL; AF185723; AAF05502.1; --

DR EMBL; AF185724; AAF05503.1; --

DR EMBL; AF185725; AAF05504.1; --

DR EMBL; AF185726; AAF05505.1; --

DR EMBL; AF185727; AAF05506.1; --

DR EMBL; AF185728; AAF05507.1; --

DR HSP; P19134; ITFD.

DR GO; GO:0005576; C:extracellular; IEA.

DR GO; GO:0008199; F:ferric iron binding; IEA.

DR GO; GO:0006879; P:iron ion homeostasis; IEA.

DR GO; GO:0006826; P:iron ion transport; IEA.

DR InterPro; IPR001156; Transferrin.

DR Pfam; PF00405; transferrin; 1.

FT NON\_TER 1 1

FT NON\_TER 36 36

SQ SEQUENCE 36 AA; 3802 MW; E2DD122186D5FA8 CRC64;

Query Match 95.0%; Score 19; DB 6; Length 36;

Best Local Similarity 100.0%; Pred. No. 5.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

|  
|  
|  
28 VAEF 31

RESULT 7

Q48420

ID Q48420 PRELIMINARY; PRT; 44 AA.

AC Q48420;

DT 01-NOV-1996 (TRENBLrel. 01, Created)

DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE Hypothetical protein (Fragment).

OS Klebsiella pneumoniae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Klebsiella.

OC NCBI\_TaxID=573;

OC [1]

SEQUENCE FROM N.A.

```

C STRAIN=ATCC 25955;
A Willard B.L.;
T "Investigation of the Klebsiella pneumoniae 1,3-propanediol pathway:
T Characterization and expression of glycerol dehydratase and 1,3-
T propanediol oxidoreductase."
L Thesis (1994), Chemical Engineering, University of Wisconsin-Madison.
N [2]
P SEQUENCE FROM N.A.
C STRAIN=ATCC 25955;
A Skraly F.A., Willard B.L., Cameron D.C.;
T "The dha regulon of Klebsiella pneumoniae."
L Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
C -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
C -I- SIMILARITY: BELONGS TO THE MIP/AQUAPORIN FAMILY (TC 1.A.8).
R EMBL; U30903; AAA74254.1; -
R HSP; P11244; 1FX8
R GO; GO:0016021; C:integral to membrane; IEA.
R GO; GO:0005215; F:transporter activity; IEA.
R GO; GO:0006810; P:transport; IEA.
R InterPro; IPR000425; MIP.
R Pfam; PF00230; MIP; 1.
R ProDom; PD000295; MIP family; 1.
M Hypothetical protein-Transmembrane; Transport.
I NON_TER 44
C SEQUENCE 44 AA; 4537 MW; 39851658FF88E734 CRC64;

Query Match 95.0%; Score 19; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/ 2 VAEF 5
C 12 VAEF 15

RESULT 8
330H9
C Q830H9 PRELIMINARY; PRT; 45 AA.
C Q830H9;
T 01-JUN-2003 (TREMBlrel. 24, Created)
T 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
T 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
E Hypothetical protein.
E EF2805.
E Enterococcus faecalis (Streptococcus faecalis).
E Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
E NCBI_TaxID=1351;
E [1]
P SEQUENCE FROM N.A.
C STRAIN=V583 / ATCC 700802;
C MEDLINE=22550857; PubMed=12663927;
A Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
A Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.P.,
A Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
A Daugherty S., DeBoy R.T., Durkin S., Kolony J., Madupu R., Nelson M.,
A Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
A Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
E "Role of mobile DNA in the evolution of vancomycin-resistant
E Enterococcus faecalis."
E Science 299:2071-2074 (2003).
E EMBL; AE016955; AA082500.1; -
E TIGR; EF2805; -
P SEQUENCE 45 AA; 5278 MW; 89CB55F5CB4C014A CRC64;

Query Match 95.0%; Score 19; DB 16; Length 45;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/ 2 VAEF 5
C 21 VAEF 24

```

```

RESULT 9
O25478 PRELIMINARY; PRT; 48 AA.
AC O25478;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein HP0789.
GN HP0789.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]_TaxID=210;
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547 (1997).
DR EMBL; AE000591; AA007847.1; -
DR PIR; B64618; E64618.
DR TIGR; HP0789; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 48 AA; 5492 MW; ED5911D96F57BFEB CRC64;

Query Match 95.0%; Score 19; DB 16; Length 48;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 15 VAEF 18

RESULT 10
Q70XL5
AC Q70XL5 PRELIMINARY; PRT; 51 AA.
AC Q70XL5;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
DE RB1255.
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]_TaxID=117;
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gieckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
DR EMBL; BX294135; CAD71991.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 51 AA; 5951 MW; 60F62389C0FDA6A7 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 51;

```

Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||||  
31 VAEF 34

SULT 11  
4941

Q44941 PRELIMINARY; PRT; 52 AA.  
Q44941; 1996 (TrEMBLrel. 01, Created)  
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
Short ORF.  
LON.  
Bacillus brevis (Brevibacillus brevis).  
Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.  
NCBI\_TaxID=1393;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=HPD31;  
ITO K., Ueda S., Yamagata H.;  
"Cloning, characterization, and inactivation of the Bacillus brevis  
lon gene";  
J. Bacteriol. 174:2281-2287(1982).  
EMBL; D00863; BAA00736.1; -;  
PIR; A42375; A42375.  
SEQUENCE 52 AA; 5558 MW; 5D689099F84F9AAF CRC64;  
Query Match 95.0%; Score 19; DB 2; Length 52;  
Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||||  
47 VAEF 50

SULT 12  
X3Y8

Q8X3Y8 PRELIMINARY; PRT; 54 AA.  
Q8X3Y8;  
01-MAR-2002 (TrEMBLrel. 20, Created)  
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
Hypothetical protein 22382.  
Z382.  
Escherichia coli O157:H7.  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.  
NCBI\_TaxID=83334;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=O157:H7 / EDL933 / ATCC 700927;  
MEDLINE=21074935; PubMed=11206551;  
Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Poramousis K.,  
Apodaca J., Anantharaman I.S., Lin J., Yen G., Schwartz D.C.,  
Welch R.A., Blattner F.R.;  
"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
Nature 409:529-533(2001).  
EMBL; AE005369; AAC56425.1; -;  
PIR; E85745; E85745.  
Hypothetical protein; Complete proteome.  
SEQUENCE 54 AA; 6507 MW; 8B5FACD98E6140CE CRC64;  
Query Match 95.0%; Score 19; DB 16; Length 54;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||||  
Db 32 VAEF 35

RESULT 13

Q8FIF3 PRELIMINARY; PRT; 54 AA.  
AC Q8FIF3;  
01-MAR-2003 (TrEMBLrel. 23, Created)  
01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
GN C1469.  
OS Escherichia coli O6.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=217992;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=06:H1 / CFT073 / ATCC 700928;  
RX MEDLINE=22388234; PubMed=12471157;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
"Extensive mosaic structure revealed by the complete genome sequence  
of uropathogenic Escherichia coli";  
Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
DR EMBL; AE016759; AAN79938.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 54 AA; 5322 MW; 6B3C3D0F21ED4376 CRC64;  
Query Match 95.0%; Score 19; DB 16; Length 54;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||||  
Db 4 VAEF 7

RESULT 14

Q8VSE8 PRELIMINARY; PRT; 61 AA.  
AC Q8VSE8;  
01-MAR-2002 (TrEMBLrel. 20, Created)  
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN CP0186.  
OS Shigella flexneri 2a.  
OC Plasmid pCP301.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Shigella.  
OX NCBI\_TaxID=42897;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=301;  
Jin O., Zhang J.Y., Liu H., Yang J., Yang F., Zhang X.B., Wang J.H.,  
Yang G.W., Wu H.T., Dong J., Sun L.L., Xue Y., Zhao A.L., Gao Y.S.,  
Zhu Z.P., Xu B., Chen S.X., Yao Z.J., He B.K., Chen R.S., Ma D.L.,  
Yuan Z.H., Xu J.G., Wang Y., Shen Y., Lu W.C., Qiang B.Q., Wen Y.M.,  
Hou Y.D.;  
"Complete DNA sequence and analysis of the large virulence plasmid  
pCP301 of Shigella flexneri";  
Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF386526; AAL72547.1; -;  
GO; GO:0046821; C:extrachromosomal DNA; IEA.  
KW Hypothetical protein; Plasmid.  
SQ SEQUENCE 61 AA; 6858 MW; F1CC17B10B28CBFC CRC64;

Query Match 95.0%; Score 19; DB 2; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
 b 54 VAEF 57

RESULT 15  
 9P166 PRELIMINARY; PRT; 61 AA.  
 C Q9P166;  
 I 01-OCT-2000 (TRENBLrel. 15, Created)  
 T 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
 T 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
 S PRO2435  
 S Homo sapiens (Human)  
 C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 X NCBI\_TaxID=9606;  
 N [1]  
 P SEQUENCE FROM N.A.  
 C TISSUE=Liver;  
 A Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,  
 A He F.;  
 I "Functional prediction of the coding sequences of 79 new genes deduced  
 T by analysis of cDNA clones from human fetal liver."  
 L Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 R EMBL; AF119881; AAF69635.1; -  
 Q SEQUENCE 61 AA; 7467 MW; 6433950D4657C1D4 CRC64;

Query Match 95.0%; Score 19; DB 4; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
 b 30 VAEF 33

RESULT 16  
 99G29 PRELIMINARY; PRT; 68 AA.  
 C Q99G29;  
 T 01-JUN-2001 (TRENBLrel. 17, Created)  
 T 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
 T 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 E OSF49 (Hypothetical protein) (Unknown)  
 S Helicoverpa zea single nucleocapsid nucleopolyhedrovirus,  
 S Helicoverpa armigera nucleopolyhedrovirus G4, and  
 S Helicoverpa armigera nuclear polyhedrosis virus.  
 C Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 C Nucleopolyhedrovirus.  
 X NCBI\_TaxID=10468, 148363, 51313;  
 N [1]  
 P SEQUENCE FROM N.A.  
 C SPECIES=Helicoverpa zea single nucleocapsid nucleopolyhedrovirus;  
 A Chen X., Zhang W.-J., Wong J., Chun G., Lu A., McCutchen B.P.,  
 A Presnail J.K., Herrmann R., Dolan M., Tingey S., Hu Z.-H., Vlask J.M.;  
 T "Genome sequence analysis of Helicoverpa zea single nucleocapsid  
 T nucleopolyhedrovirus."  
 T Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 L [2]  
 P SEQUENCE FROM N.A.  
 C SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;  
 A Deng F., Chen X., Vlask J.M., Arif B.M., Hu Z.;  
 T "Sequence analysis of the gp37 gene of Helicoverpa armigera single-  
 T nucleocapsid nucleopolyhedrovirus."  
 T Zhongguo Bingdaxue 15:35-42(2000).  
 L [3]  
 P SEQUENCE FROM N.A.  
 C SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;

RA Wang H., Hu Z., Sun X., Vlask J.M., Chen X.;  
 RT "Sequence analysis of the iap3 gene of Helicoverpa armigera single-  
 RT nucleocapsid nucleopolyhedrovirus."  
 RL Zhongguo Bingdaxue 15:43-49(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;  
 RX MEDLINE=21078302; PubMed=11210934;  
 RA Wang H., Chen X., Wang H., Arif B.M., Vlask J.M., Hu Z.;  
 RT "Nucleotide sequence and transcriptional analysis of a putative basic  
 RT DNA-binding protein of Helicoverpa armigera polyhedrovirus."  
 RL Virus Genes 22:113-120(2001).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;  
 RX MEDLINE=21064569; PubMed=1125177;  
 RA Chen X., Ijkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,  
 RA Peters S., Zuidema D., Lankhorst R.K., Vlask J.M., Hu Z.;  
 RT "The sequence of the Helicoverpa armigera single-nucleocapsid  
 RT nucleopolyhedrovirus genome."  
 RL J. Gen. Virol. 82:241-257(2001).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;  
 RA Chen X., Ijkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,  
 RA Peters S., Zuidema D., Lankhorst R.K., Vlask J.M., Hu Z.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Helicoverpa armigera nuclear polyhedrosis virus; STRAIN=C1;  
 RX PubMed=12050807;  
 RA Zhang C.X., Wu J.C.;  
 RT "Genome structure and the p10 gene of the Helicoverpa armigera  
 RT nucleopolyhedrovirus."  
 RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 33:179-184(2001).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Helicoverpa armigera nuclear polyhedrosis virus; STRAIN=C1;  
 RA Zhang C.X., Jin W.R.;  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Helicoverpa armigera nuclear polyhedrosis virus;  
 RA Fang M., Hu Z., Chen X., Vlask J.M.;  
 RT "Genetic organization of the HindIII-L region of Helicoverpa armigera  
 RT single-nucleocapsid nucleopolyhedrovirus."  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF334030; AAL56194.1; -  
 DR EMBL; AF271059; AAG53791.1; -  
 DR EMBL; AF303045; AAK36298.1; -  
 DR EMBL; AF266694; AAK64316.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 68 AA; 7962 MW; 61B7718BFBB195FF CRC64;

Query Match 95.0%; Score 19; DB 12; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAEF 5  
 Db 29 VAEF 32

RESULT 17  
 Q82X12 PRELIMINARY; PRT; 68 AA.  
 ID Q82X12;  
 AC Q82X12;  
 DT 01-JUN-2003 (TRENBLrel. 24, Created)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Helix-turn-helix protein, CopG family.  
 GN NE0289.  
 OS Nitrosomonas europaea.

Bacteria; Proteobacteria; Betaproteobacteria; Nitrospomonadales;  
Nitrospomonadaceae; Nitrospomonas.  
NCBI\_TaxID=915;  
[1]

SEQUENCE FROM N.A.  
STRAIN=ATCC 19718 / IF0 14298;  
MEDLINE=22586410; PubMed=12700255;  
Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,  
Hauser L., Hooper A.B., Klotz M.G., Norton J., Savavedra-Soto L.A.,  
Acierno D.M., Hommes N.G., Whittaker M.M., Arp D.J.;  
"Complete genome sequence of the ammonia-oxidizing bacterium and  
obligate chemolithoautotroph Nitrospomonas europaea.";  
J. Bacteriol. 185:2759-2773(2003).  
EMBL; BX321857; CAD84200.1; -;  
Complete proteome.  
SEQUENCE 68 AA; 7800 MW; 8A8338537524DD6 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 69;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||||  
31 VAEF 34

SULT 18  
C7L1

QX7L1 PRELIMINARY; PRT; 69 AA.

01-OCT-2003 (TrEMBLrel. 25, Created)  
01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
OSJNBa004M1.9.22 protein (OSJNBa0053B21.1 protein).  
OSJNBa004M1.9.22 OR OSJNBa0053B21.1.  
Oryza sativa (Rice)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzae; Oryza.  
NCBI\_TaxID=4530;  
[1]

SEQUENCE FROM N.A.  
Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,  
Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,  
Lu Y.Q., Yu S.H., Liu X.H., Qian Y.J., Lu Y., Li C., Li T.,  
Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,  
Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,  
Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,  
Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.W., Sun Y.,  
Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,  
Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;  
Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
EMBL; AL731601; CAE05035.1; -;  
EMBL; AL731599; CAE05527.1; -;  
SEQUENCE 69 AA; 7267 MW; 7783933BC873F0BE CRC64;

Query Match 95.0%; Score 19; DB 10; Length 69;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||||  
31 VAEF 34

SULT 19  
NPA7

Q8NPA7 PRELIMINARY; PRT; 69 AA.

01-OCT-2002 (TrEMBLrel. 22, Created)  
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
Hypothetical protein Cg11907.

CGL1907.  
OS Corynebacterium glutamicum (Brevibacterium flavum).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=1718;  
RN [1]  
RC SEQUENCE FROM N.A.  
RP STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
RA Nakagawa S.;  
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF005280; BAB9300.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 69 AA; 7815 MW; C0E2A072C2295DD2 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 69;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||||  
Db 23 VAEF 26

RESULT 20

Q8AUQ2 PRELIMINARY; PRT; 71 AA.

AC Q8AUQ2;  
DT 01-VAR-2003 (TrEMBLrel. 23, Created)  
DT 01-VAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Transferrin (Fragment).  
OS Salmo trutta (Brown trout).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.  
OX NCBI\_TaxID=8032;  
RN [1]

SEQUENCE FROM N.A.  
RP STRAIN=Str-1, Str-2, Str-3, Str-4, Str-9, Str-10, and Str-11;  
RX MEDLINE=22135992; PubMed=12140239;  
RA Antunes A., Templeton A.R., Guyomard R., Alexandrino P.;  
RT "The role of nuclear genes in intraspecific evolutionary inference:  
genealogy of the transferrin gene in the brown trout.";  
Mol. Biol. Evol. 19:1272-1287(2002).  
DR EMBL; AF48850; ANI17027.1; -;  
DR EMBL; AF48849; ANI17027.1; JOINED.  
DR EMBL; AF48858; ANI17032.1; -;  
DR EMBL; AF48857; ANI17032.1; JOINED.  
DR EMBL; AF48866; ANI17037.1; -;  
DR EMBL; AF48865; ANI17037.1; JOINED.  
DR EMBL; AF48874; ANI17042.1; -;  
DR EMBL; AF48873; ANI17042.1; JOINED.  
DR EMBL; AF488914; ANI17068.1; -;  
DR EMBL; AF488913; ANI17068.1; JOINED.  
DR EMBL; AF488921; ANI17073.1; -;  
DR EMBL; AF488930; ANI17078.1; -;  
DR EMBL; AF488929; ANI17078.1; JOINED.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0008199; F:iron ion binding; IEA.  
DR GO; GO:0006879; P:iron ion homeostasis; IEA.  
DR GO; GO:0006826; P:iron ion transport; IEA.  
DR InterPro; IPR001156; Transferrin.  
DR Pfam; PF00405; transferrin; 1.  
DR PRINTS; PR00422; TRANSFERRIN.

FT NON\_TER 1 1  
FT NON\_TER 71 71  
SQ SEQUENCE 71 AA; 7546 MW; 892E38F22DD2AC4D CRC64;

Query Match 95.0%; Score 19; DB 13; Length 71;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Y      2 VAEF 5
b      33 VAEF 36

ESULT 21
BAUP0 PRELIMINARY; PRT; 71 AA.
C Q8AUP0;
T 01-MAR-2003 (TrEMBLrel. 23, Created)
T 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
T 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
E Transferrin (Fragment).
S Salmo salar (Atlantic salmon).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Actinopterygii; Neopterygii; Teleostei; Euteleostei;
C Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
C NCBI_TaxID=8030;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=Ssa-1, and Ssa-2;
X MEDLINE=22135932; PubMed=12140239;
A Antunes A., Templeton A.R., Guyomard R., Alexandrino P.;
I "The role of nuclear genes in intraspecific evolutionary inference:
I Genealogy of the transferrin gene in the brown trout.";
L Mol. Biol. Evol. 19:1272-1287(2002).
R EMBL; AF488834; AAN17017.1; -.
R EMBL; AF488833; AAN17017.1; JOINED.
R EMBL; AF488842; AAN17022.1; -.
R EMBL; AF488841; AAN17022.1; JOINED.
R GO; GO:000576; C:extracellular; IEA.
R GO; GO:0008199; P:ferrous iron binding; IEA.
R GO; GO:0006879; P:iron ion homeostasis; IEA.
R GO; GO:0006826; P:iron ion transport; IEA.
R Pfam; PF00405; Transferrin.
R PRINTS; PR00422; TRANSFERRIN.
I NON_TER 1
I NON_TER 71 71
Q SEQUENCE 71 AA; 7504 MW; 9D86A1234CBACAD CRC64;

Query Match 95.0%; Score 19; DB 13; Length 71;
Best Local Similarity 100.0%; Pred. No. 1.le+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y      2 VAEF 5
b      33 VAEF 36

ESULT 22
97252 PRELIMINARY; PRT; 72 AA.
C P97252;
T 01-MAY-1997 (TrEMBLrel. 03, Created)
T 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
T 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
E Late control gene D protein (Fragment).
N D.
S Escherichia coli.
C Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
C Enterobacteriaceae; Escherichia.
C NCBI_TaxID=562;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=KL2;
X MEDLINE=97251358; PubMed=9097040;
A Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
A Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
A Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
A Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
A Sivasubram S., Tagami H., Takeda J., Takemoto K., Wada C.,

Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
DR EMBL; D90847; BAA15949.1; -.
DR EMBL; D90846; BAA15938.1; -.
FT NON_TER 1
SQ SEQUENCE 72 AA; 8099 MW; 42A63B25B00EADCB CRC64;

Query Match 95.0%; Score 19; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.le+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
Db      5 VAEF 8

RESULT 24
Q82MX0 PRELIMINARY; PRT; 74 AA.
ID Q82MX0;
AC Q82MX0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN SAV1532.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
```



2A Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C., T.,  
2B Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,  
2C Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
2D "Genome sequence of an industrial microorganism Streptomyces  
2E avermitilis: deducing the ability of producing secondary  
2F metabolites";  
2G Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
2H [2]  
2I SEQUENCE FROM N.A.  
2J STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
2K MEDLINE=22608306; PubMed=12692562;  
2L Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
2M Sakaki Y., Hattori M., Omura S.;  
2N "Complete genome sequence and comparative analysis of the industrial  
2O microorganism Streptomyces avermitilis";  
2P Nat. Biotechnol. 21:526-531(2003).  
2Q EMBL; AP005027; BAC69243.1; -  
2R Hypothetical protein; Complete proteome.  
2S SEQUENCE 74 AA; 7960 MW; D8CE71C8B0A8B01 CRC64;  
2T  
2U Query Match 95.0%; Score 19; DB 16; Length 74;  
2V Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;  
2W Matches 4; Conservative 0; Mismatches 0;  
2X 2 VAEF 5  
2Y |||||  
2Z 36 VAEF 39

RESULT 25  
2862M8 PRELIMINARY; PRT; 77 AA.  
29 ID Q862M8  
30 AC Q862M8  
31 DT 01-JUN-2003 (TRENBLrel. 24, Created)  
32 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
33 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
34 DE Similar to MAD2 protein (Fragment).  
35 OS Bos taurus (Bovine).  
36 DE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
37 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;  
38 OC Bovidae; Bovinae; Bos.  
39 NCBI\_TaxID=9913;  
40 [1]  
41 SEQUENCE FROM N.A.  
42 RX MEDLINE=22544902; PubMed=12658628;  
43 RA Ishikawa H., Katsuma S., Kizaki K., Patel O.V., Nakano H.,  
44 Takahashi T., Imai K., Hirasawa A., Shiojima S., Ikawa H., Suzuki Y.,  
45 Tsujimoto G., Izaike Y., Todoroki J., Hashizume K.;  
46 "Characterization of gene expression profiles in early bovine  
47 pregnancy using a custom cDNA microarray";  
48 RL Mol. Reprod. Dev. 65:9-18(2003).  
49 DR EMBL; AB098952; BAC56442.1; -  
50 DR InterPro; IPR003511; DNABind\_HORMA.  
51 DR PROSITE; PS50815; HORMA; 1.  
52 FT NON\_TER 1 77  
53 FT NON\_TER 77 77  
54 SQ SEQUENCE 77 AA; 8832 MW; E2E1D625C948E571 CRC64;  
55

Query Match 95.0%; Score 19; DB 6; Length 77;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0;  
2 VAEF 5  
6 VAEF 9

RESULT 26  
Q88VS7 PRELIMINARY; PRT; 77 AA.  
ID Q88VS7  
AC Q88VS7  
DT 01-JUN-2003 (TRENBLrel. 24, Created)  
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Hypothetical protein.  
GN LP 1960.  
OS Lactobacillus plantarum.  
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
OC Lactobacillus.  
OX NCBI\_TaxID=1590;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCIMB 8826 / WCFS1;  
RX MEDLINE=22480296; PubMed=12566566;  
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,  
RA Kuipers O.P., Leer R., Turchini R., Peters S.A., Sarda-brink H.M.,  
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,  
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,  
RA De Vos W.M., Siezen R.J.;  
RA "Complete genome sequence of Lactobacillus plantarum WCFS1";  
RT Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).  
RL EMBL; AL935257; CAD64344.1; -  
DR Hypothetical protein; Complete proteome.  
KW SEQUENCE 77 AA; 8618 MW; C00E2BFF1D401F2F CRC64;  
SQ  
35 Query Match 95.0%; Score 19; DB 16; Length 77;  
36 Best Local Similarity 100.0%; Pred. No. 1.2e+03; Indels 0; Gaps 0;  
37 Matches 4; Conservative 0; Mismatches 0;  
38 2 VAEF 5  
39 |||||  
40 68 VAEF 71

RESULT 27  
Q98582 PRELIMINARY; PRT; 79 AA.  
ID Q98582  
AC Q98582  
DT 01-FEB-1997 (TRENBLrel. 02, Created)  
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE A532L protein.  
GN A532L.  
OS Paramyxium bursaria chlorella virus 1 (PBCV-1).  
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.  
OX NCBI\_TaxID=10506;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96400190; PubMed=8806566;  
RA Kutish G.P., Li Y., Lu Z., Furuta M., Rock D.L., Van Etten J.L.;  
RA "Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome: map  
RA positions 182 to 258";  
RT Virology 223:303-317(1996).  
RL [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20013326; PubMed=10544099;  
RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., King W.,  
RA Lisee A.D., Nickerson K.W., Van Etten J.L.;  
RA "Chlorella virus PBCV-1 encodes a functional homospemidine  
RA synthase";  
RL Virology 263:254-262(1999).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20478054; PubMed=11021991;  
RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;  
RA "Characterization of a beta-1,3-glucanase encoded by chlorella virus  
RA PBCV-1";  
RL Virology 276:27-36(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Van Etten J.L.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Van Etten J.L.;  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

RN [6] SEQUENCE FROM N.A.  
 RP Van Eten J.L.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Graves M.V., Van Eten J.L.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RA Graves M.V., Van Eten J.L.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RA Gurnon J.R., Graves M.V., Van Eten J.L.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U42580; AAC96899.1; -.  
 DR PIR; T18034; T18034.  
 SQ SEQUENCE 79 AA; 8698 MW; B191C627F5D5C5A7 CRC64;

Query Match 95.0%; Score 19; DB 12; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 Db 49 VAEF 52

RESULT 28  
 Q9KI36 PRELIMINARY; PRT; 80 AA.  
 AC Q9KI36;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Yag.  
 GN isg.  
 OS Agrobacterium tumefaciens.  
 OG Plasmid Ti.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 CX NCBI\_TaxID=358;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Winans S.C., Zhu J., Oger P.M., Schrammeyer B., Hooykaas P.J.,  
 RA Farrand S.K.;  
 RT "Octopine-type Ti plasmid sequence."  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF242881; AAF77180.1; -.  
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
 KW Plasmid.  
 SQ SEQUENCE 80 AA; 8723 MW; 4B9379C1D907CABE CRC64;

Query Match 95.0%; Score 19; DB 2; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 Db 65 VAEF 68

RESULT 29  
 Q89WL6 PRELIMINARY; PRT; 81 AA.  
 AC Q89WL6;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Bar0662 protein.  
 GN BSR0662.  
 OS Bradyrhizobium japonicum.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Bradyrhizobium.  
 CX NCBI\_TaxID=375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=USDA 110;  
 RX MEDLINE=22484998; PubMed=12597275;  
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
 RA Sasamoto S., Watanabe A., Ideawa K., Iiguchi M., Kawashima K.,  
 RA Kohara N., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,  
 RA Tabata S.;  
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
 RT Bradyrhizobium japonicum USDA110."  
 RL DNA Res. 9:189-197(2002).  
 DR EMBL; AP005937; BAC45927.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 81 AA; 9235 MW; D512A9FFED0DA6C7 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 Db 43 VAEF 46

RESULT 30  
 Q9KAV4 PRELIMINARY; PRT; 84 AA.  
 AC Q9KAV4;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein BH2182.  
 GN BH2182.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 CX NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis."  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 DR EMBL; AP001514; BAB05901.1; -.  
 DR PIR; F83922; F83922.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 84 AA; 9669 MW; 6CDE3768ED9F5D84 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 Db 38 VAEF 41

RESULT 31  
 Q821D2 PRELIMINARY; PRT; 84 AA.  
 AC Q821D2;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN CCA01010.  
 OS Chlamydomophila caviae.

```
XC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
XX NCBI_TaxID=83557;
XN [1]
XP SEQUENCE FROM N.A.
RC STRAIN=GPIC;
XA MEDLINE=22569155; PubMed=12682364;
YA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
ZA Heidelberg J., Holtzapple E., Khouri H., Federova N.B., Carty H.A.,
ZA Umayam L.A., Haft D.H., Peterson J., Beanan M.J., White O.,
ZA Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G., Bavol P.M.,
ZA Fraser C.M.;
XT "Genome sequence of Chlamydophila caviae (Chlamydia psittaci GPIC):
XT examining the role of niche-specific genes in the evolution of the
XT Chlamydiaceae.";
XL Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL; AE016997; AAP05749.1; -.
DR TIGR; CCA01010; -.
GW Hypothetical protein; Complete proteome.
SQ SEQUENCE 84 AA; 9191 MW; 7DF609729C7093A CRC64;

Query Match 95.0%; Score 19; DB 16; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ZY 2 VAEF 5
Zb 68 VAEF 71

RESULT 32
ZYVPV6 PRELIMINARY; PRT; 84 AA.
AC Q7VPV6;
JT 01-OCT-2003 (TrEMBLrel. 25, Created)
JT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
JT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
ZE Hypothetical protein.
ZN CP80775.
XS Chlamydia pneumoniae (Chlamydophila pneumoniae).
XC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
XN NCBI_TaxID=83558;
XN [1]
XP SEQUENCE FROM N.A.
RC STRAIN=TW-193;
XA Geng M.M., Schuhmacher A., Muehldorfer I., Bersch K.W., Schaefer K.P.,
XA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
XT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
XT other Chlamydia strains based on whole genome sequence analysis.";
XL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
XR EMBL; AE017159; AAP98704.1; -.
ZW Hypothetical protein.
SQ SEQUENCE 84 AA; 9260 MW; 7786DF8BDD7B3AB8 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ZY 2 VAEF 5
Zb 68 VAEF 71

RESULT 33
ZY3116 PRELIMINARY; PRT; 85 AA.
AC O43116;
JT 01-JUN-1998 (TrEMBLrel. 06, Created)
JT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
JT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
ZE Hypothetical protein (fragment).
XS Colletotrichum gloeosporioides (Anthracocone fungus) (Glomerella
XS cingulata).
XC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
```

```
OC Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
OC Glomerella.
XX NCBI_TaxID=5457;
XN [1]
XP SEQUENCE FROM N.A.
RC STRAIN=UO62;
YA Stephenson S.-A., Maclean D.J., Manners J.M.;
YA "Disruption of a novel pathogenicity gene of Colletotrichum
YA gloeosporioides results in a hypersensitive response in the host
YA Stylosanthes guianensis.";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U94183; AAB92222.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001387; HTH_3.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SMO0530; HTH_XRE; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 85 AA; 9316 MW; BFB9A0E5F44B9CF2 CRC64;

Query Match 95.0%; Score 19; DB 3; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ZY 2 VAEF 5
Zb 45 VAEF 48

RESULT 34
ZY9149 PRELIMINARY; PRT; 87 AA.
AC Q99149;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transferrin (fragments).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
XX NCBI_TaxID=9940;
XN [1]
XP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=91177867; PubMed=1849850;
XA Tu G.F., Achen M.G., Aldred A.R., Southwell B.R., Schreiber G.;
RT "The distribution of cerebral expression of the transferrin gene is
RT species specific.";
RL J. Biol. Chem. 266:6201-6208(1991).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
DR EMBL; M64691; AAA31585.1; -.
DR EMBL; M64692; AAA31586.1; -.
DR PIR; A38725; A38725.
DR HSSP; P19134; ITFD.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0006879; P:iron ion homeostasis; IEA.
DR GO; GO:0006826; P:iron ion transport; IEA.
DR InterPro; IPR001156; Transferrin.
DR SMART; SMO0094; TR_FER; 1.
DR PROSITE; PS00205; TRANSFERRIN_1; 1.
KW Iron transport; Metal-binding.
FT NON_TER 1
FT NON_TER 44
FT NON_TER 87
SQ SEQUENCE 87 AA; 9433 MW; AA464B2ABDC92FAB CRC64;

Query Match 95.0%; Score 19; DB 6; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      2 VAEF 5
Db      13 VAEF 16

RESULT 35
QY7Y3Y9 PRELIMINARY; PRT; 87 AA.
AC QY7Y3Y9;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Bacteriophage PY54.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
CX NCBI_TaxID=172667;
RN [1]
RP SEQUENCE FROM N.A.
RA Hertwig S.; Klein I., Schmidt V., Beck S., Hammerl J.A., Appel B.;
RT "Sequence analysis of the genome of the temperate Yersinia
RT enterocolitica phage PY54."
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Hertwig S.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ564013; CAD91786.1; -.
KW Exonuclease; Hypothetical protein.
SQ SEQUENCE 87 AA; 8538 MW; D17D7A7E3075459B CRC64;

Query Match 95.0%; Score 19; DB 9; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
Db      31 VAEF 34

RESULT 36
Q9S590 PRELIMINARY; PRT; 88 AA.
AC Q9S590;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE Cytochrome P460 (Fragment).
GN CYP.
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
CX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RA Iizumi T., Nakamura K.;
RT "Construction of tryptophan requiring-mutant of Nitrosomonas europaea
RT by inactivation of trpC gene using homologous recombination."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS030031; BA883388.1; -.
FT NON-TER 88
SQ SEQUENCE 88 AA; 9765 MW; A59003348F4E280C CRC64;

Query Match 95.0%; Score 19; DB 2; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
Db      29 VAEF 32

RESULT 37
Q41185 PRELIMINARY; PRT; 88 AA.
AC Q41185;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Glyceraldehyde-3-phosphate-dehydrogenase subunit GapB (Fragment);
GN GAPB.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
CX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93013005; PubMed=1398114;
RA Shih M.C.; Heinrich P.; Goodman H.M.;
RT "Cloning and chromosomal mapping of nuclear genes encoding chloroplast
RT and cytosolic glyceraldehyde-3-phosphate-dehydrogenase from
RT Arabidopsis thaliana."
RL Gene 119:317-319(1992).
DR EMBL; S45911; AAB23533.1; -.
GO; GO:0009507; Chloroplast; IEA.
KW Chloroplast.
FT NON-TER 88
SQ SEQUENCE 88 AA; 9198 MW; FCDB6A7E37B6999D CRC64;

Query Match 95.0%; Score 19; DB 10; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
Db      38 VAEF 41

RESULT 38
Q9ZGY2 PRELIMINARY; PRT; 89 AA.
AC Q9ZGY2;
DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN Yil03.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
CX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=KIM10+;
RC MEDLINE=99043898; PubMed=9826348;
RA Lindler L.R.; Plano G.V.; Burland V.; Mayhew G.F.; Blattner F.R.;
RT "Complete DNA sequence and detailed analysis of the Yersinia pestis
RT KIM5 plasmid encoding murine toxin and capsular antigen."
RL Infect. Immun. 66:5731-5742(1998).
DR EMBL; AF074611; AAC82761.1; -.
DR FIR; T15018; T15018.
DR GO; GO:0046821; Extrachromosomal DNA; IEA.
DR InterPro; IPR003006; IG MHC.
DR PROSITE; PS00290; IG MHC; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 89 AA; 10736 MW; 72B4221126023EE4 CRC64;

Query Match 95.0%; Score 19; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5

```

Query Match 95.0%; Score 19; DB 16; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 3 VAEF 6

RESULT 39

ID Q9CNX1 PRELIMINARY; PRT; 89 AA.

AC Q9CNX1

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE RPS15.

DE RPS15 OR PMO301.

RN Pasteurella multocida.

SC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Pasteurella.

CX NCBI\_TaxID=747;

RN RN

SEQUENCE FROM N.A.

RC STRAIN=Pm70;

RC MEDLINE=211145866; PubMed=11248100;

RA May B.J., Zhang O., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;

RT "Complete genomic sequence of Pasteurella multocida Pm70";

RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

DR EMBL; AF006066; AAK02385.1; -.

DR HSP; PF0378; IAB3.

DR GO; GO:0005622; C:intracellular; IEA.

DR GO; GO:0005840; C:ribosome; IEA.

DR GO; GO:0003735; F:structural constituent of ribosome; IEA.

DR GO; GO:0006412; P:protein biosynthesis; IEA.

DR InterPro; IPR000589; Ribosomal\_S15.

DR InterPro; IPR005280; Ribosomal\_S15\_b.

DR Pfam; PF00312; Ribosomal\_S15; 1.

DR ProDom; PD157043; RS15\_bact; 1.

DR TIGRFAMs; TIGR00952; S15\_bact; 1.

DR PROSITE; PS00362; RIBOSOMAL\_S15; 1.

KW Complete proteome.

SQ SEQUENCE 89 AA; 10184 MW; F796E0AB5283ED0E CRC64;

Query Match 95.0%; Score 19; DB 16; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 12 VAEF 15

RESULT 40

ID Q9CFD1 PRELIMINARY; PRT; 89 AA.

AC Q9CFD1

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Hypothetical protein YPJB.

GN YPJB OR LL1550.

OC Lactococcus lactis (subsp. lactis) (Streptococcus lactis).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

CX NCBI\_TaxID=1360;

RN RN

SEQUENCE FROM N.A.

RC STRAIN=IL1403;

RC MEDLINE=21235186; PubMed=11337471;

RA Bolotin A., Winkler P., Mauger S., Jaillon O., Malarme K.,

RA Weissenbach J., Ehrlich S.D., Sorokin A.;

RT "The complete genome sequence of the lactic acid bacterium Lactococcus

RT lactis sp. lactis IL1403.";

RL Genome Res. 11:731-753(2001).

DR EMBL; AF006385; AAK05648.1; -.

DR PIR; F86818; F86818.

DR KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 89 AA; 10403 MW; D90DF3CF71D12D06 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2 VAEF 5

RESULT 41

ID Q8IZX3 PRELIMINARY; PRT; 90 AA.

AC Q8IZX3

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DR MAD2 mitotic arrest deficient-like 1 variant.

GN MAD2L1.

CS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

CX NCBI\_TaxID=9606;

RN [1]

SEQUENCE FROM N.A.

RA Yin F., Fan D.M.;

RT "Identifying a new variant of MAD2L1.";

RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AF394735; AAN74648.1; -.

DR InterPro; IPR003511; DNAbind\_HORMA.

DR Pfam; PF02301; HORMA; 1.

DR PROSITE; PS50815; HORMA; 1.

SQ SEQUENCE 90 AA; 10335 MW; 8209F5A7AD8D09B CRC64;

Query Match 95.0%; Score 19; DB 4; Length 90;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 20 VAEF 23

RESULT 42

ID Q98TB2 PRELIMINARY; PRT; 91 AA.

AC Q98TB2

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Preproinsulin (fragment).

OS Ambloplites rupestris (Rock bass).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;

OC Centrarchidae; Ambloplites.

CX NCBI\_TaxID=109273;

RN [1]

SEQUENCE FROM N.A.

RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;

RT "Molecular cloning of preproinsulin cDNA from the rock bass.";

RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.

CC -|- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

CC EMBL; AF195884; AAK28708.1; -.

DR HSP; PF1308; ILPH.

DR GO; GO:0005576; C:extracellular; IEA.

DR GO; GO:0005179; F:hormone activity; IEA.

DR GO; GO:0007582; P:physiological processes; IEA.

DR InterPro; IPR004825; Ins/IGF/relax.

DR Pfam; PF00049; Insulin; 1.

DR PRINTS; PR00277; INSULINE.

DR SWART; SMO0078; IGF; 1.

Query Match 95.0%; Score 19; DB 16; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 3 VAEF 6

RESULT 39  
Q9CNX1 PRELIMINARY; PRT; 89 AA.  
ID Q9CNX1  
AC Q9CNX1  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE RPS15.  
EN RPS15 OR PMO301.  
NCBI\_TaxID=747;  
RN RNP15  
RS PASTEURILLA MULTOCIDA.  
SC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
SS Pasteurellaceae; Pasteurella.  
CX NCBI\_TaxID=747;  
GN PASTEURELLA  
LN PASTEURELLA  
RC STRAIN=Pm70;  
PC MEDLINE=211145866; PubMed=11248100;  
RA MAY B.J., Zhang O., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida Pm70";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
EMBL; AF006066; AAK02385.1; -;  
DR HSP; P80378; IAB3.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0005840; C:ribosome; IEA.  
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
DR GO; GO:0006412; P:protein biosynthesis; IEA.  
DR InterPro; IPR000589; Ribosomal\_S15.  
DR InterPro; IPR005280; Ribosomal\_S15\_b.  
DR Pfam; PF00312; Ribosomal\_S15; 1.  
DR ProDom; PD157043; RS15\_bact; 1.  
DR TIGRFAMs; TIGR00952; S15\_bact; 1.  
DR PROSITE; PS00362; RIBOSOMAL\_S15; 1.  
KW Complete proteome.  
SQ SEQUENCE 89 AA; 10184 MW; F796E0AB5283ED0E CRC64;

Query Match 95.0%; Score 19; DB 16; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 12 VAEF 15

RESULT 40  
Q9CFD1 PRELIMINARY; PRT; 89 AA.  
ID Q9CFD1  
AC Q9CFD1  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein YPJB.  
GN YPJB OR LL1550.  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
OX NCBI\_TaxID=1360;  
RN [1]  
RS SEQUENCE FROM N.A.  
RC STRAIN=IL1403;  
PC MEDLINE=21235186; PubMed=11337471;  
RA Bolotin A., Winkler P., Mauger S., Jaillon O., Malarme K.,  
Weissenbach J., Ehrlich S.D., Sorokin A.;  
RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
lactis ssp. lactis IL1403".  
RL Genome Res. 11:731-753(2001).  
EMBL; AF006385; AAK05648.1; -;  
DR PIR; F86818; F86818.  
DR KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 89 AA; 10403 MW; D90DF3CF71D12D06 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2 VAEF 5

RESULT 41  
Q8IZX3 PRELIMINARY; PRT; 90 AA.  
ID Q8IZX3  
AC Q8IZX3  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE MAD2 mitotic arrest deficient-like 1 variant.  
GN MAD2L1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
CX NCBI\_TaxID=9606;  
RN [1]  
RS SEQUENCE FROM N.A.  
RY Yin F., Fan D.M.;  
RA "Identifying a new variant of MAD2L1";  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF394735; AAN74648.1; -;  
DR InterPro; IPR003511; DNAbind\_HORMA.  
DR Pfam; PF02301; HORMA; 1.  
DR PROSITE; PS50815; HORMA; 1.  
SQ SEQUENCE 90 AA; 10335 MW; 8209F5A7AD8D09B CRC64;

Query Match 95.0%; Score 19; DB 4; Length 90;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 20 VAEF 23

RESULT 42  
Q98TB2 PRELIMINARY; PRT; 91 AA.  
ID Q98TB2  
AC Q98TB2  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Preproinsulin (fragment).  
OS Ambloplites rupestris (Rock bass).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;  
OC Centrarchidae; Ambloplites.  
OX NCBI\_TaxID=109273;  
RN [1]  
RS SEQUENCE FROM N.A.  
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;  
RT "Molecular cloning of preproinsulin cDNA from the rock bass.";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
CC -|- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.  
CC -|- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.  
DR EMBL; AF199584; AAK28708.1; -;  
DR HSP; P01308; IZPH.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0005179; F:hormone activity; IEA.  
DR GO; GO:0007582; P:physiological processes; IEA.  
DR InterPro; IPR004825; Ins/IGF/relax.  
DR Pfam; PF00049; Insulin; 1.  
DR PRINTS; PR00277; INSULINE.  
DR SWART; SMO0078; IGF; 1.

Query Match 95.0%; Score 19; DB 16; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2 VAEF 5

RESULT 43  
Q9CNX1 PRELIMINARY; PRT; 89 AA.  
ID Q9CNX1  
AC Q9CNX1  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE RPS15.  
EN RPS15 OR PMO301.  
NCBI\_TaxID=747;  
RN RNP15  
RS PASTEURILLA MULTOCIDA.  
SC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
SS Pasteurellaceae; Pasteurella.  
CX NCBI\_TaxID=747;  
GN PASTEURELLA  
LN PASTEURELLA  
RC STRAIN=Pm70;  
PC MEDLINE=211145866; PubMed=11248100;  
RA MAY B.J., Zhang O., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida Pm70";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
EMBL; AF006066; AAK02385.1; -;  
DR HSP; P80378; IAB3.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0005840; C:ribosome; IEA.  
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
DR GO; GO:0006412; P:protein biosynthesis; IEA.  
DR InterPro; IPR000589; Ribosomal\_S15.  
DR InterPro; IPR005280; Ribosomal\_S15\_b.  
DR Pfam; PF00312; Ribosomal\_S15; 1.  
DR ProDom; PD157043; RS15\_bact; 1.  
DR TIGRFAMs; TIGR00952; S15\_bact; 1.  
DR PROSITE; PS00362; RIBOSOMAL\_S15; 1.  
KW Complete proteome.  
SQ SEQUENCE 89 AA; 10184 MW; F796E0AB5283ED0E CRC64;

Query Match 95.0%; Score 19; DB 16; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 12 VAEF 15

RESULT 44  
Q9CFD1 PRELIMINARY; PRT; 89 AA.  
ID Q9CFD1  
AC Q9CFD1  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein YPJB.  
GN YPJB OR LL1550.  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
OX NCBI\_TaxID=1360;  
RN [1]  
RS SEQUENCE FROM N.A.  
RC STRAIN=IL1403;  
PC MEDLINE=21235186; PubMed=11337471;  
RA Bolotin A., Winkler P., Mauger S., Jaillon O., Malarme K.,  
Weissenbach J., Ehrlich S.D., Sorokin A.;  
RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
lactis ssp. lactis IL1403".  
RL Genome Res. 11:731-753(2001).  
EMBL; AF006385; AAK05648.1; -;  
DR PIR; F86818; F86818.  
DR KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 89 AA; 10403 MW; D90DF3CF71D12D06 CRC64;

```

DR PROSITE; PS00262; INSULIN; 1.
FT NON TER 1 1
FT NON TER 91 91
SQ SEQUENCE 91 AA; 10100 MW; B86CB8256DC69D39 CRC64;

Query Match
Best Local Similarity 95.0%; Score 19; DB 13; Length 91;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
Db 55 VAEF 58

RESULT 43
Q8KV12 PRELIMINARY; PRT; 94 AA.
AC Q8KV12
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative transcriptional regulator (Fragment).
OS Rhizobium etli.
CG Plasmid p42b.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
CX NCBI_TaxID=29449;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CE3;
PA Cevallos M.A., Izquierdo J., Porta H., Tun C., Davila G., Brom S.;
RT "Rhizobium etli CE3 contains at least three plasmids of the RepABC
family: A structural and an evolutionary analysis.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF313446; AN88940.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
FT NON TER 1 1
SQ SEQUENCE 94 AA; 10222 MW; C09D6EA3DFF1E2BA CRC64;

Query Match
Best Local Similarity 95.0%; Score 19; DB 2; Length 94;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
Db 38 VAEF 41

RESULT 44
O46425 PRELIMINARY; PRT; 94 AA.
AC O46425
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aquaporin 1 (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
CX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand;
PA Carter E.P., Umenishi F., Matthey M.A., Verkman A.S.;
RT "Increased water permeability across the blood-gas barrier in rabbit
lungs in the first 24 hours after birth.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SURCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC EMBL; AF000311; AAB94408.1; -.
DR HSP; P29972; 1FOY.
DR GO; GO:0016021; C:integral to membrane; IEA.

```

---

```

DR GO; GO:005741; C:mitochondrial outer membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000425; MIP.
DR Pfam; PF00230; MIP; 1.
DR ProDom; PD000295; MIP_family; 1.
DR PROSITE; PS00221; MIP; 1.
KW Porin; Transmembrane; Transport.
FT NON TER 1 1
FT NON TER 94 94
SQ SEQUENCE 94 AA; 10057 MW; 963D5527631E8CDC CRC64;

Query Match
Best Local Similarity 95.0%; Score 19; DB 6; Length 94;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
Db 9 VAEF 12

RESULT 45
Q8U8Q1 PRELIMINARY; PRT; 96 AA.
AC Q8U8Q1
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Atu4039.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
CX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Moo L.,
RA Chen Y., Paulsen I.T., Eissen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeter E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58.";
RL Science 294:2317-2323 (2001).
DR EMBL; AE009334; AAL44940.1; -.
DR PIR; AB3053; AB3053.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 96 AA; 11193 MW; FBD635894B46A8AE CRC64;

Query Match
Best Local Similarity 95.0%; Score 19; DB 16; Length 96;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
Db 55 VAEF 58

RESULT 46
Q89WX7 PRELIMINARY; PRT; 96 AA.
AC Q89WX7
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bs10551 protein.

```

```

IN BSL0551.
NS Bradyrhizobium japonicum.
NC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
XC Bradyrhizobiaceae; Bradyrhizobium.
XX NCBI_TaxID=375;
XP [1]
IP SEQUENCE FROM N.A.
IC STRAIN=USDA 110;
IX MEDLINE=2484998; PubMed=12597275;
JA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
SA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
FA Kohara M., Matsumoto M., Shingo S., Tsuruoka H., Wada T., Yamada M.,
LA Tabata S.;
TT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
NT Bradyrhizobium japonicum USDA110."
IL DNA Res. 9:189-197(2002).
IR EMBL; AF005936; BAC45816.1; -.
RR GO; GO:0016020; C:membrane; IEA.
RW Pfam; PF02325; YGGT; 1.
XX Complete proteome.
XQ SEQUENCE 96 AA; 10976 MW; A9E11F9AA4BC7734 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YI 2 VAEF 5
YJ 44 VAEF 47

RESULT 47
XU0G9 PRELIMINARY; PRT; 96 AA.
IC Q8U0G9;
IT 01-JUN-2002 (TrEMBLrel. 21, Created)
YT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
YU 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
YE Hypothetical protein PF1620.
YN PF1620.
YS Pyrococcus furiosus.
XC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
XX Pyrococcus.
XN NCBI_TaxID=2261;
XP [1]
IP SEQUENCE FROM N.A.
IC STRAIN=Vc1 / DSM 3638 / ATCC 43597 / JCM 8422;
JA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
SA "The complete sequence of the Pyrococcus furiosus genome."
FA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
LA EMBL; AB010262; AAU81744.1; -.
RW Hypothetical protein; Complete proteome.
XQ SEQUENCE 96 AA; 10804 MW; D6DAE09D096D577A CRC64;

Query Match 95.0%; Score 19; DB 17; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YI 2 VAEF 5
YJ 24 VAEF 27

RESULT 48
XD Q86535 PRELIMINARY; PRT; 97 AA.
IC Q86535;
YT 01-NOV-1996 (TrEMBLrel. 01, Created)
YU 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
YE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
YN 2C/3A (Fragment).

```

```

OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LSH/S;
RX MEDLINE=92348853; PubMed=1668326;
RA Fineschi N., Cavalieri F., Garelick H., Prugnola A., Pellegrini V.,
RT Zuckerman A.J.;
RT "Characterization of a hepatitis A virus strain suitable for vaccine
production."
RL J. Hepatol. 13:S146-S151(1991).
DR EMBL; S44109; AAB22740.2; -.
FT NON_TER 1 1
FT NON_TER 97 97
SQ SEQUENCE 97 AA; 10614 MW; D207F7CB93110DDF CRC64;

Query Match 95.0%; Score 19; DB 12; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 44 VAEF 47

RESULT 49
Q8Y476 PRELIMINARY; PRT; 97 AA.
AC Q8Y476;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DB Hypothetical protein lmo2579.
GN LMO2579.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Dommann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkut G.,
RA Madueno E., Maitourman A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).
RD EMBL; AL591983; CAD00657.1; -.
DR PIR; AC1397; AC1397.
DR Listlist; LMO2579; -.
DR InterPro; IPR007138; ABM.
DR Pfam; PF03992; ABM; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 97 AA; 10979 MW; 9E758586E94218E0 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 15 VAEF 18

```

```
RESULT 50
Q9H238 Q9H238 PRELIMINARY; PRT; 99 AA.
AC Q9H238;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PA3202.
GN PA3202.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Latbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL; AE004744; AAC06590.1; -.
DR PIR; E83244; E83244.
DR InterPro; IPR005545; YCII.
DR Pfam; PF03795; YCII; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 99 AA; 10609 MW; 3AD945F4D4D5A85C CRC64;

Query Match 95.0%; Score 19; DB 16; Length 99;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 62 VAEF 65

RESULT 51
Q885M3 Q885M3 PRELIMINARY; PRT; 99 AA.
AC Q885M3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Conserved hypothetical protein.
GN P3PT01808.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Pedorova N., Tran B., Russell D.,
RA Berry K., Uterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsten T.,
RA White O., Fraser C., Collier A.;
RT "Complete sequence of Pseudomonas syringae.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016862; AAC055328.1; -.
DR TIGR; P3PT01808; -.
DR InterPro; IPR005545; YCII.
DR Pfam; PF03795; YCII; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 99 AA; 10504 MW; 62836007E4849392 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 99;

QY 2 VAEF 5
DB 62 VAEF 65

RESULT 52
Q82UR2 Q82UR2 PRELIMINARY; PRT; 100 AA.
AC Q82UR2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN NE1419.
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 / IFO 14298;
RX MEDLINE=22586410; PubMed=12700255;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
RT obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773 (2003).
DR EMBL; BX321861; CAD85330.1; -.
DR InterPro; IPR005545; YCII.
DR Pfam; PF03795; YCII; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 100 AA; 10936 MW; 82DC1153BABFAE27 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 62 VAEF 65

RESULT 53
Q8VS54 Q8VS54 PRELIMINARY; PRT; 101 AA.
AC Q8VS54;
DT 01-VAR-2002 (TrEMBLrel. 20, Created)
DT 01-VAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Elongation factor TS (Fragment).
GN TSP.
OS Lactobacillus reuteri.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 55739;
RA Nam S.J., Kim J.K., Park J.Y., Ha Y.L., Kim J.H.;
RT "Cloning of UMP-kinase gene from Lactobacillus reuteri ATCC 55739.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF401482; AAL60142.1; -.
DR GO; GO:003746; P:translational elongation factor activity; IEA.
DR GO; GO:006414; P:translational elongation; IEA.
DR InterPro; IPR001816; EF_TS.
DR Pfam; PF00889; EF_TS; 1.
DR NON_TER 1
FT SEQUENCE 101 AA; 11650 MW; B8CF30D941DB2B9B CRC64;
```



Query Match 95.0%; Score 19; DB 2; Length 101;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
 ||||  
 62 VAEF 65

RESULT 54  
 977Y2C PRELIMINARY; PRT; 101 AA.  
 C Q97YC2  
 T 01-OCT-2001 (TrEMBLrel. 18, Created)  
 T 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 T 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein SSO1404.  
 N SSO1404.  
 S Sulfolobus solfataricus.  
 X Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 X Sulfolobus.  
 X NCBI\_TaxID=2287;  
 X [1]  
 X SEQUENCE FROM N.A.  
 X STRAIN=ATCC 35092 / DSM 1617 / P2;  
 X MEDLINE=21332296; PubMed=11427726;  
 A She Q., Singh R.K., Confalonieri P., Zivanovic Y., Allard G.,  
 A Aweez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,  
 A De Moors A., Traus G., Fletcher C., Gordon P.M.K.,  
 A Heikamp-de Jong I., Jeffries A.C., Korera C.J., Medina N., Peng X.,  
 A Thi-Ngoc H.P., Redder P., Schenck M.E., Theriault C., Tolstrup N.,  
 A Garlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
 A Charrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;  
 A "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";  
 L Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
 R EMBL: AE006755; AAK41639.1;  
 R PIR: H90297; H90297.  
 R InterPro: IPR003799; DUF196.  
 R Pfam: PF02647; DUF196; 1.  
 R TIGRFAMs: TIGR01573; cas2; 1.  
 W Hypothetical protein; Complete proteome.  
 IQ SEQUENCE 101 AA; 11936 MW; A7C338AD76202E17 CRC64;

Query Match 95.0%; Score 19; DB 17; Length 101;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
 ||||  
 20 VAEF 23

RESULT 55  
 981R27 PRELIMINARY; PRT; 102 AA.  
 C Q81R27  
 T 01-JUN-2003 (TrEMBLrel. 24, Created)  
 T 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 T 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 N BA2232.  
 S Bacillus anthracis (strain Ames).  
 X Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 X NCBI\_TaxID=198094;  
 X [1]  
 X SEQUENCE FROM N.A.  
 X MEDLINE=22608414; PubMed=12721629;  
 A Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,  
 A Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,  
 A Holtzapple E.K., Ostad O.A., Helgason E., Ristone J., Wu M.,  
 A Kolonay J.F., Beaman M.J., Dodson R.J., Brinkac L.M., Gwinn M.,  
 A Debey R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,  
 A Nelson W.C., Peterson J.D., Pop M., Kouri H.M., Radune D.,

RA Banton J.L., Mahmoud Y., Jiang L., Hance I.R., Weidman J.F.,  
 RA Barry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nieman W.C.,  
 RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,  
 RA Thomson B., Friedlander A.M., Koshler T.M., Hanna P.C., Koisto A.-B.,  
 RA Fraser C.M.;  
 RA "The genome sequence of Bacillus anthracis Ames and comparison to  
 RT closely related bacteria.";  
 RL Nature 423:81-86(2003).  
 DR EMBL: AE017031; AAP26109.1; -.  
 DR TIGR; BA2232; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 102 AA; 11925 MW; 13B8E296C9A341B1 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 ||||  
 54 VAEF 57

RESULT 56  
 Q7Z2N2 PRELIMINARY; PRT; 103 AA.  
 AC Q7Z2N2  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE BK150C2.6 (Putative novel protein similar to APOBEC1 (Apolipoprotein B  
 DE mRNA editing protein) and Phorbol (Fragment).  
 GN BK150C2.6.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Matthews L.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL022318; CAB45276.1; -.  
 KW Lipoprotein.  
 FT NON\_TER 1 103  
 FT NON\_TER 103 103  
 SQ SEQUENCE 103 AA; 12146 MW; 5DC969AE3ED348D9 CRC64;

Query Match 95.0%; Score 19; DB 4; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 ||||  
 46 VAEF 49

RESULT 57  
 Q9F368 PRELIMINARY; PRT; 103 AA.  
 ID Q9F368  
 AC Q9F368  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein SCO4420.  
 GN SCO4420 OR SC6F11.18.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 CX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21956410; PubMed=12000953;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hdaigo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinovitch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,  
 RA Hopwood D.A.,  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RL coelicolor A3(2).",  
 RL Nature 417:141-147(2002).  
 DR EMBL; AL939120; CAC08429.1; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 103 AA; 11387 MW; 55C2B80589EB75B5 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 Db 54 VAEF 57

RESULT 58  
 Q89MS9 PRELIMINARY; PRT; 103 AA.  
 ID Q89MS9  
 AC Q89MS9  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DE BL4113 protein.  
 DE BL4113 protein.  
 GN BLR4113.  
 OS Bradyrhizobium japonicum.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Bradyrhizobium.  
 OX NCBI\_TaxID=375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=USDA 110;  
 RC MEDLINE=22484998; PubMed=12597725;  
 RA Kaneko T., Nakamura Y., Sato S., Minamiasawa K., Uchiyama T.,  
 RA Sasamoto S., Watanabe A., Iidesawa K., Iiguchi M., Kawashina K.,  
 RA Kohara M., Matsumoto M., Shimpo S., Tsuruguchi H., Wada T., Yamada M.,  
 RA Tabata S.;  
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
 RT Bradyrhizobium japonicum USDA110.",  
 RL DNA Res. 9:189-197(2002).  
 DR EMBL; AP005950; BAC49378.1; --  
 KW Complete proteome.  
 SQ SEQUENCE 103 AA; 11104 MW; BF27CB9F90FB5723 CRC64;

Query Match 95.0%; Score 19; DB 16; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 Db 96 VAEF 99

RESULT 59  
 Q856B5 PRELIMINARY; PRT; 104 AA.  
 ID Q856B5  
 AC Q856B5  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE GP57.  
 OS Mycobacteriophage Barnyard.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
 OX NCBI\_TaxID=205860;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=22592660; PubMed=12705866;  
 RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,  
 RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,  
 RA Bruckner W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,  
 RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,  
 RA Hatfull G.F.;  
 RT "Origins of highly mosaic mycobacteriophage genomes.",  
 RL Cell 113:171-182(2003).  
 RL EMBL; AY129339; AAN02111.1; --  
 DR EMBL; AY129339; AAN02111.1; --  
 SQ SEQUENCE 104 AA; 11074 MW; BB806EAC401B9FD0 CRC64;

Query Match 95.0%; Score 19; DB 9; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 Db 59 VAEF 62

RESULT 60  
 Q8SVK0 PRELIMINARY; PRT; 105 AA.  
 ID Q8SVK0  
 AC Q8SVK0  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical protein ECU05\_0770.  
 GN ECU05\_0770.  
 OS Encephalitozoon cuniculi.  
 OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.  
 OX NCBI\_TaxID=6035;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-M1;  
 RC Genoscope;  
 RC Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-M1;  
 RX MEDLINE=21576510; PubMed=11719806;  
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,  
 RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,  
 RA Delbac F., El Alsaoui H., Peyret F., Saurin W., Gouy M.,  
 RA Weissenbach J., Vivares C.P.;  
 RT "Genome sequence and gene compaction of the eukaryote parasite  
 RT Encephalitozoon cuniculi.",  
 RL Nature 414:450-453(2001).  
 DR EMBL; AL590445; CAD26596.1; --  
 KW Hypothetical protein.  
 SQ SEQUENCE 105 AA; 11897 MW; 08059108C05D3BB9 CRC64;

Query Match 95.0%; Score 19; DB 5; Length 105;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 Db 77 VAEF 80

RESULT 61  
 Q856U1 PRELIMINARY; PRT; 105 AA.  
 ID Q856U1  
 AC Q856U1  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE GP3.  
 OS Mycobacteriophage Corndog.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
 OX NCBI\_TaxID=205875;

3N SEQUENCE FROM N.A.  
3P MEDLINE=22592660; PubMed=12705866;  
3A Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,  
3A Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,  
3A Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,  
3A Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,  
3A Hatfull G.F.,  
3R "Origins of highly mosaic mycobacteriophage genomes.";  
3T Cell 113:171-182(2003).  
3R EMBL; AY129335; AAN01935.1; -  
3Q SEQUENCE 105 AA; 11955 MW; 7D0EC09F5F2AF2DD CRC64;  
  
Query Match 95.0%; Score 19; DB 9; Length 105;  
Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
2Y 2 VAEF 5  
2b 41 VAEF 44  
  
RESULT 62  
ID Q96Z31 PRELIMINARY; PRT; 105 AA.  
AC Q96Z31;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein ST2000.  
EN ST2000.  
CS Sulfolobus tokodaii.  
DC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
DD Sulfolobus.  
DX NCBI\_TaxID=111955;  
XN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=JCM 10545 / 7;  
XZ MEDLINE=21456156; PubMed=11572479;  
ZA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
ZA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,  
ZA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
ZA Yoshizawa T., Tanaka T., Kudon Y., Yamazaki J., Kishida N., Oguchi A.,  
ZA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
ZA Oshima T., Kikuchi H.;  
RT "Complete genome sequence of an aerobic thermoacidophilic  
RT Crenarchaeon, Sulfolobus tokodaii strain7.";  
RL DNA Res. 8:123-140(2001).  
JR EMBL; AP000988; BAB67055.1; -  
JR InterPro; IPR006783; Transposase\_29.  
JR Pfam; PF04693; Transposase\_29; 2.  
GW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 105 AA; 11997 MW; AFE6ABD4DF2DECOE CRC64;  
  
Query Match 95.0%; Score 19; DB 17; Length 105;  
Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
2Y 2 VAEF 5  
2b 29 VAEF 32  
  
RESULT 63  
ID P71530 PRELIMINARY; PRT; 106 AA.  
AC P71530;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE NIFX.  
DS Methanococcus maripaludis.

OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanococcaceae; Methanococcus.  
OX NCBI\_TaxID=39152;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LL;  
RX MEDLINE=96011360; PubMed=7592322;  
RA Blank C.E., Kessler P.S., Leigh J.A.;  
RT "Genetics in methanogens: transposon insertion mutagenesis of a  
RT Methanococcus maripaludis nifH gene.";  
RL J. Bacteriol. 177:5773-5777(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LL;  
RX MEDLINE=97144542; PubMed=8990309;  
RA Kessler P.S., McIarnan J., Leigh J.A.;  
RT "Nitrogenase phylogeny and the molybdenum dependence of nitrogen  
RT fixation in Methanococcus maripaludis.";  
RL J. Bacteriol. 179:541-543(1997).  
DR EMBL; U75887; AAC45519.1; -  
DR PIR; T10097; T10097.  
DR InterPro; IPR003731; DUF153.  
DR Pfam; PF02579; NitroPeMo-Co; 1.  
SQ SEQUENCE 106 AA; 11862 MW; 4734D89CFED33F3B CRC64;  
  
Query Match 95.0%; Score 19; DB 1; Length 106;  
Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 VAEF 5  
Db 35 VAEF 38  
  
RESULT 64  
Q9BGL5 PRELIMINARY; PRT; 106 AA.  
ID Q9BGL5;  
AC Q9BGL5;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Matrilysin (Fragment).  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Butheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Endometrium;  
RA Smith G.W., Rieke W.A., Caesar C.A., Smith M.F.;  
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF267158; AAG59846.1; -  
DR HSSP; P09237; IMMQ.  
DR MEROPS; M10.008; -  
DR GO; GO:0005578; C:extracellular matrix; IEA.  
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR006026; Peptidase\_M.  
DR InterPro; IPR001818; Pept\_M10A\_M12B.  
DR Pfam; PF00413; Peptidase\_M10; 1.  
DR PRINTS; PR00138; MATRIXIN.  
DR SMART; SM00235; ZnMc; 1.  
FT NON\_TER 1  
FT NON\_TER 106  
FT NON\_TER 106  
SQ SEQUENCE 106 AA; 11578 MW; EBD271054928018 CRC64;  
  
Query Match 95.0%; Score 19; DB 6; Length 106;  
Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 VAEF 5  
Db 35 VAEF 38

```

Db      6 VAEF 9
Query Match      95.0%; Score 19; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 65
O23682
ID      O23682      PRELIMINARY;      PRT;      106 AA.
AC
DT      01-JAN-1998 (TrEMBLrel. 05, Created)
DT      01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT      01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE      Cathepsin B-like cysteine proteinase (Fragment).
GN      T7123.13.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=cv. Columbia;
RA      Federstepfel N.A., Conway A.B., Conway A.R., Davis K., Brendel V.,
RA      Palm C.J., Au M., Araujo R., Chung E., Kurtz D.B., Buehler E.,
RA      Dewar K., Feng J., Kim C., Li Y., Shinn P., Sun H., Oji O.,
RA      Osborne B., Shen Y.K., Toriumi M., Vyotskaia V., Yu G., Theologis A.,
RA      Ecker J., Davis R.W.;
RT      "Genomic sequence of Arabidopsis BAC T7123."
RL      Submitted (JUN-1998) to the EMBL/GenBank/DBSJ databases.
DR      EMBL; U89595; AAC24377.1; -.
FT      NON TER      106
FT      NON TER      106
SQ      SEQUENCE      106 AA; 11743 MW; C5D5EA5897AD17AP CRC64;

Query Match      95.0%; Score 19; DB 10; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
Db      75 VAEF 78

RESULT 66
O9C1L1
ID      O9C1L1      PRELIMINARY;      PRT;      107 AA.
AC
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Putative thioredoxin G6G8.7.
OS      Neurospora crassa.
OC      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC      Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX      NCBI_TaxID=5141;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=74-OR23-1A;
RA      MEDLINE=21135672; PubMed=11238395;
RA      Bean L.E., Dvorachek W.H. Jr., Braun E.L., Errett A., Saenz G.S.,
RA      Giles M.D., Werner-Washburne M., Nelson M.A., Natvig D.O.;
RT      "Analysis of the ptx-1 (snz-1/sno-1) region of the Neurospora crassa
RT      genome. Correlation of pyridoxine-requiring phenotypes with mutations
RT      in two structural genes."
RL      Genetics 157:1067-1075 (2001).
DR      EMBL; AF309689; AAK07845.1; -.
FT      HSP; P80028; 1TQF
DR      GO; GO:0005489; P:electron transporter activity; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR006662; ThioRed.
DR      InterPro; IPR006663; ThioRedox_dom2.
DR      Pfam; PF00085; ThioRed.
DR      PRINTS; PR00421; THIOREDOXIN.
DR      Redox-active center.
KW      Redox-active center.
SQ      SEQUENCE      107 AA; 11676 MW; CF456EAF85BE3776 CRC64;

Query Match      95.0%; Score 19; DB 5; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
Db      40 VAEF 43

RESULT 67
O8IHV3
ID      O8IHV3      PRELIMINARY;      PRT;      107 AA.
AC
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Hypothetical protein.
GN      PF11_0423
OS      Plasmodium falciparum (isolate 3D7).
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=36329;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=3D7;
RA      MEDLINE=22255705; PubMed=12368864;
RA      Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA      Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA      Risen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA      Chant M.-S., Nene V., Shallow J.J., Suh B., Peterson J., Anguoli S.,
RA      Perlea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA      Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA      McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA      Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA      Fraser C.M., Barrett B.;
RT      "Genome sequence of the human malaria parasite Plasmodium
RT      falciparum."
RL      Nature 419:498-511 (2002).
DR      EMBL; AE014842; RAN36006.1; -.
KN      Hypothetical protein.
SQ      SEQUENCE      107 AA; 12714 MW; CA351258FCF46E35 CRC64;

Query Match      95.0%; Score 19; DB 5; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
Db      40 VAEF 43

RESULT 68
O8NRX6
ID      O8NRX6      PRELIMINARY;      PRT;      107 AA.
AC
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Uncharacterized ACR.
GN      CGL0915.
OS      Corynebacterium glutamicum (Brevibacterium flavum).
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX      NCBI_TaxID=1718;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA      Nakagawa S.;
RT      "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RL      Submitted (MAY-2002) to the EMBL/GenBank/DBSJ databases.
DR      EMBL; AP005276; BAB98308.1; -.
DR      InterPro; IPR007138; ABM.
DR      Pfam; PF03992; ABM; 1.
DR      Complete proteome.
KW      Complete proteome.

```

IQ SEQUENCE 107 AA; 12373 MW; B5ACF23621078018 CRC64;  
 Query Match 95.0%; Score 19; DB 16; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YV 2 VAEF 5  
 |||||  
 22 VAEF 25

RESULT 69  
 9PUB82 PRELIMINARY; PRT; 109 AA.  
 AC O9PUB2;  
 UC O9PUB2; 13, Created)  
 YT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 YT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 YT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Retinol binding protein (fragment).  
 JS Crocodylus niloticus (Nile crocodile) (African crocodile).  
 XC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 XC Archosauria; Crocodylia; Crocodylinae; Crocodylus.  
 XX NCBI\_TaxID=8501;  
 XN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX Hughes S., Zelus D., Mouchiroud D.;  
 RA "Warm-blooded isochore structure in Nile crocodile and turtle.";  
 RL Mol. Biol. Evol. 16:1521-1527(1999).  
 DR EMBL; AJ011392; CAB56418.1; -.  
 DR HSSP; P02753; IRBP.  
 DR GO; GO:0005215; P:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR002345; Lipocalin.  
 DR InterPro; IPR000566; Lipocln\_cycPABP.  
 DR Pfam; PF00061; lipocalin; 1.  
 DR PRINTS; PR00179; LIPOCALIN.  
 DR NON\_TER 1  
 FT NON TER 1  
 FT NON TER 109 109  
 SQ SEQUENCE 109 AA; 12552 MW; 0DCAE93895B91A23 CRC64;

Query Match 95.0%; Score 19; DB 13; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 |||||  
 25 VAEF 28

RESULT 70  
 OS3333 PRELIMINARY; PRT; 109 AA.  
 ID OS3333;  
 AC OS3333;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein RV3183.  
 GN RV3183 OR MTV014.27 OR RV3275.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL021646; CAA16648.1; -.  
 DR EMBL; AE007140; AAK47615.1; -.  
 DR FTR; E70949; E70949.  
 DR TIGR; MT275; -.  
 DR TubercuList; RV3183; -.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR InterPro; IPR001387; HTH\_3.  
 DR Pfam; PF01381; HTH\_3; 1.  
 DR SMART; SMO0530; HTH\_XRE; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 109 AA; 11802 MW; A66DBE160D79613A CRC64;

Query Match 95.0%; Score 19; DB 16; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 |||||  
 97 VAEF 100

Db 97 VAEF 100

RESULT 71  
 Q7TX28 PRELIMINARY; PRT; 109 AA.  
 ID Q7TX28;  
 AC Q7TX28;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Possible transcriptional regulatory protein.  
 GN MB3209.  
 OS Mycobacterium bovis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1765;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AF2122/97;  
 RX MEDLINE=22709107; PubMed=12788972;  
 RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,  
 RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsepe C., Simon S.,  
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
 RT "The complete genome sequence of Mycobacterium bovis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).  
 DR EMBL; BX248345; CAB95301.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 109 AA; 11802 MW; A66DBE160D79613A CRC64;

Query Match 95.0%; Score 19; DB 16; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 |||||  
 97 VAEF 100

Db 97 VAEF 100

```

RESULT 72
ID OI1343 PRELIMINARY; PRT; 110 AA.
AC OI1343;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE XI-5 protein (Fragment).
GN XI-5.
OS Molluscum contagiosum virus subtype 1 (MCV1).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Molluscipoxvirus.
OX NCBI_TaxID=10280;
RN [1]
SEQUENCE FROM N.A.
RA Moratilla M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,
RA Lopez-Estebaranz J.L., Esteban M., Martin-Gallardo A.;
RT "A Random DNA Sequencing, Computer-Based Approach for the Generation
RT of a Gene Map of Molluscum Contagiosum Virus.";
RL Virus Genes 0:0-0(1997)
DR EMBL; U86919; AB57977.1; -.
FT NON_TER 1
FT NON_TER 110
SQ SEQUENCE 110 AA; 11856 MW; 6AE9CE25FC3DB780 CRC64;

Query Match 95.0%; Score 19; DB 12; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 9 VAEF 12

RESULT 73
Q88EG3 PRELIMINARY; PRT; 110 AA.
AC Q88EG3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE YCII-related domain protein.
GN PF4502.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=22423080; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,
RA Uterback T., Rizzio M., Lee K., Kosack D., Noestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hohneskel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016790; AAN70076.1; -.
DR TIGR; PF4502; -.
DR InterPro; IPR005545; YCII.
DR Pfam; PF03795; YCII; 1.
KW Complete proteome.
SQ SEQUENCE 110 AA; 11775 MW; 774B31C0D925A01E CRC64;

Query Match 95.0%; Score 19; DB 16; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 VAEF 5
DB 73 VAEF 76

RESULT 74
Q8U296 PRELIMINARY; PRT; 111 AA.
AC Q8U296;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein PF0943.
GN PF0943.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010208; AAL81067.1; -.
DR InterPro; IPR007842; HEPN.
DR Pfam; PF05168; DUF712; 1.
DR PROSITE; PS0910; HSEN; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 111 AA; 13379 MW; C6A005AE2B6CA4E CRC64;

Query Match 95.0%; Score 19; DB 17; Length 111;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 48 VAEF 51

RESULT 75
Q86PL2 PRELIMINARY; PRT; 112 AA.
AC Q86PL2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nuclear receptor NHR-1 (Fragment).
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RA Robinson-Rechavi M., Maina C.V., Gissendanner C.R., Laudet V.,
RA Sluder A.;
RT "Explosive lineage-specific expansion of the orphan nuclear receptor
RT HNF4 in nematodes.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY204162; AAO39166.1; -.
DR GO; GO:0004872; Fireceptor activity; IEA.
DR InterPro; IPR008946; Str_ncl_receptor.
KW Receptor.
FT NON_TER 1
FT NON_TER 112
SQ SEQUENCE 112 AA; 12826 MW; 44FC60AE3744791 CRC64;

Query Match 95.0%; Score 19; DB 5; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 4 VAEF 51

```

b 13 VAEF 16

earch completed: May 24, 2004, 17:41:03  
ob time : 123.714 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

M protein - protein search, using sw model

run on: May 24, 2004, 17:38:13 ; Search time 15 Seconds  
(without alignments)  
17.209 Million cell updates/sec

file: US-09-594-978A-1

effect score: 20

sequence: 1 XVAEF 5

coring table: BLOSUM62

Gapol 10.0 , Gapext 0.5

sarched: 389414 segs, 51625971 residues

total number of hits satisfying chosen parameters: 389414

imum DB seq length: 0

aximum DB seq length: 200000000

set-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

atabase :

Issued Patents AA: \*  
1: /cgn2\_6/prodata/2/iaa/5A-COMB.pep: \*  
2: /cgn2\_6/prodata/2/iaa/5B-COMB.pep: \*  
3: /cgn2\_6/prodata/2/iaa/5A-COMB.pep: \*  
4: /cgn2\_6/prodata/2/iaa/5B-COMB.pep: \*  
5: /cgn2\_6/prodata/2/iaa/PCTUS-COMB.pep: \*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	19	95.0	7	4	US-09-724-566A-78
2	19	95.0	8	4	US-09-724-566A-81
3	19	95.0	9	4	US-08-197-484-86
4	19	95.0	9	4	US-08-197-484-145
5	19	95.0	9	4	US-09-724-566A-73
6	19	95.0	9	5	PCT-US95-02121-86
7	19	95.0	9	5	PCT-US95-02121-145
8	19	95.0	14	4	US-09-724-566A-72
9	19	95.0	14	4	US-09-724-566A-97
10	19	95.0	28	2	US-08-733-825-3
11	19	95.0	35	1	US-08-487-890A-69
12	19	95.0	35	1	US-08-468-763-1
13	19	95.0	35	2	US-08-393-996A-1
14	19	95.0	35	2	US-08-478-435-69
15	19	95.0	35	2	US-08-337-483-69
16	19	95.0	35	2	US-08-478-373-69
17	19	95.0	35	3	US-08-474-671-69
18	19	95.0	35	3	US-08-483-577A-69
19	19	95.0	35	3	US-08-897-438-69
20	19	95.0	35	4	US-08-637-654-69
21	19	95.0	35	4	US-08-849-518-69
22	19	95.0	68	4	US-09-393-634-80
23	19	95.0	69	4	US-09-489-039A-13555
24	19	95.0	71	4	US-09-328-352-7111
25	19	95.0	73	1	US-08-321-071A-10
26	19	95.0	74	4	US-08-894-139-10
27	19	95.0	88	4	US-09-732-210-1412

19	95.0	93	1	US-07-612-674-12	Sequence 12, Appl
19	95.0	98	4	US-09-543-681A-5816	Sequence 5816, Ap
19	95.0	100	2	US-08-464-517-8	Sequence 8, Appli
19	95.0	100	2	US-08-246-361A-8	Sequence 8, Appli
19	95.0	100	2	US-08-463-772-8	Sequence 8, Appli
19	95.0	100	5	PCT-US93-05000-8	Sequence 21, Appl
19	95.0	101	2	US-08-580-988A-21	Sequence 3612, Ap
19	95.0	103	4	US-09-540-236-3612	Sequence 25, Appl
19	95.0	106	2	US-08-464-517-25	Sequence 25, Appl
19	95.0	106	2	US-08-246-361A-25	Sequence 25, Appl
19	95.0	106	3	US-08-463-772-25	Sequence 8721, Ap
19	95.0	106	4	US-09-489-039A-8721	Sequence 25, Appl
19	95.0	106	5	PCT-US93-05000-25	Sequence 9315, Ap
19	95.0	115	4	US-09-489-039A-9915	Sequence 21614, A
19	95.0	122	4	US-09-252-991A-21614	Sequence 7200, Ap
19	95.0	123	4	US-09-489-039A-7200	Sequence 53, Appl
19	95.0	129	4	US-08-529-055-53	Sequence 12835, A
19	95.0	143	4	US-09-489-039A-12835	Sequence 167, App
19	95.0	145	4	US-09-198-452A-167	Sequence 4292, Ap
19	95.0	150	2	US-09-134-000C-4292	Sequence 3, Appli
19	95.0	150	3	US-08-460-694-3	Sequence 3, Appli
19	95.0	150	3	US-08-460-744-3	Sequence 3, Appli
19	95.0	150	3	US-07-667-711B-3	Sequence 3, Appli
19	95.0	152	4	US-08-679-493A-208	Sequence 208, App
19	95.0	153	4	US-09-134-000C-4087	Sequence 4087, Ap
19	95.0	156	4	US-09-252-991A-20805	Sequence 20805, A
19	95.0	164	4	US-09-370-838-203	Sequence 203, App
19	95.0	166	4	US-09-540-236-2222	Sequence 2222, Ap
19	95.0	172	1	US-08-471-058-16	Sequence 16, Appl
19	95.0	172	1	US-08-690-095-3	Sequence 3, Appli
19	95.0	172	3	US-08-471-057-16	Sequence 16, Appl
19	95.0	172	3	US-09-113-789-3	Sequence 3, Appli
19	95.0	172	4	US-08-470-865-16	Sequence 16, Appl
19	95.0	173	1	US-08-193-977-4	Sequence 4, Appli
19	95.0	174	4	US-09-489-039A-9869	Sequence 9869, Ap
19	95.0	175	2	US-08-737-980-2	Sequence 2, Appli
19	95.0	180	4	US-09-328-352-6739	Sequence 6739, Ap
19	95.0	185	4	US-09-252-991A-24341	Sequence 24341, A
19	95.0	189	4	US-09-543-681A-7313	Sequence 7313, Ap
19	95.0	190	1	US-08-815-241-1	Sequence 1, Appli
19	95.0	190	3	US-09-128-385-1	Sequence 1, Appli
19	95.0	194	4	US-09-134-000C-3800	Sequence 3800, Ap
19	95.0	199	3	US-09-232-200-85	Sequence 85, Appl
19	95.0	199	4	US-09-232-197-85	Sequence 85, Appl
19	95.0	199	4	US-09-232-201-85	Sequence 85, Appl
19	95.0	199	4	US-09-232-195-85	Sequence 85, Appl
19	95.0	203	4	US-09-107-532A-4612	Sequence 4612, Ap
19	95.0	205	2	US-08-684-024-1	Sequence 1, Appli
19	95.0	205	2	US-08-684-024-6	Sequence 6, Appli
19	95.0	205	2	US-08-684-024-7	Sequence 7, Appli
19	95.0	205	3	US-09-145-868-1	Sequence 1, Appli
19	95.0	205	3	US-09-145-868-6	Sequence 6, Appli
19	95.0	205	3	US-09-145-868-7	Sequence 7, Appli
19	95.0	208	4	US-09-134-000C-3408	Sequence 3408, Ap
19	95.0	212	4	US-09-489-039A-10557	Sequence 10557, A
19	95.0	213	3	US-09-232-191-17	Sequence 17, Appl
19	95.0	213	3	US-09-232-200-17	Sequence 17, Appl
19	95.0	213	4	US-09-232-201-17	Sequence 17, Appl
19	95.0	213	4	US-09-232-197-17	Sequence 17, Appl
19	95.0	213	4	US-09-232-195-17	Sequence 17, Appl
19	95.0	222	4	US-09-252-991A-30586	Sequence 30586, A
19	95.0	222	4	US-09-328-352-8105	Sequence 8105, Ap
19	95.0	225	4	US-09-489-847-155	Sequence 155, App
19	95.0	226	4	US-09-328-352-7363	Sequence 7363, Ap
19	95.0	228	4	US-09-455-960-8	Sequence 8, Appli
19	95.0	231	4	US-09-328-352-7924	Sequence 7924, Ap
19	95.0	231	4	US-09-134-000C-6021	Sequence 6021, Ap
19	95.0	232	4	US-09-489-039A-8897	Sequence 8897, Ap
19	95.0	234	3	US-09-248-335-34	Sequence 34, Appl
19	95.0	234	3	US-09-107-532A-6523	Sequence 6523, Ap
19	95.0	236	4	US-09-605-858-34	Sequence 34, Appl
19	95.0	238	4	US-09-605-858-35	Sequence 35, Appl
19	95.0	238	4	US-09-598-401C-68	Sequence 68, Appl



101	19	95.0	241	4	US-09-134-001C-5598	Sequence 5598, Ap	174	19	95.0	337	4	US-09-371-671B-2	Sequence 2, Appli
102	19	95.0	241	4	US-09-489-039A-7795	Sequence 7795, Ap	175	19	95.0	338	4	US-09-328-352-7151	Sequence 7151, Ap
103	19	95.0	243	4	US-09-134-001C-3587	Sequence 3587, Ap	176	19	95.0	339	3	US-09-345-468-3	Sequence 3, Appli
104	19	95.0	246	4	US-09-252-991A-30976	Sequence 30976, A	177	19	95.0	339	4	US-09-414-453A-3	Sequence 3, Appli
105	19	95.0	247	4	US-09-230-136-16	Sequence 16, Appl	178	19	95.0	339	4	US-09-134-000C-5209	Sequence 5209, Ap
106	19	95.0	247	4	US-09-372-422A-48	Sequence 48, Appl	179	19	95.0	342	4	US-09-543-681A-4207	Sequence 4207, Ap
107	19	95.0	249	3	US-09-154-083-2	Sequence 2, Appli	180	19	95.0	345	4	US-09-107-532A-3849	Sequence 3849, Ap
108	19	95.0	249	3	US-09-345-468-9	Sequence 9, Appli	181	19	95.0	346	4	US-09-286-365-107	Sequence 107, App
109	19	95.0	249	4	US-09-372-422A-22	Sequence 22, Appl	182	19	95.0	347	4	US-09-134-001C-3298	Sequence 3298, Ap
110	19	95.0	249	4	US-09-414-453A-9	Sequence 9, Appli	183	19	95.0	350	4	US-09-655-270A-17	Sequence 17, Appl
111	19	95.0	249	4	US-09-252-991A-22610	Sequence 22610, A	184	19	95.0	350	4	US-09-651-941-21	Sequence 21, Appl
112	19	95.0	250	1	US-08-234-939-2	Sequence 2, Appli	185	19	95.0	350	4	US-09-955-597-21	Sequence 21, Appl
113	19	95.0	250	1	US-08-558-865-2	Sequence 2, Appli	186	19	95.0	350	4	US-09-252-991A-27528	Sequence 27528, A
114	19	95.0	250	3	US-08-654-025-2	Sequence 2, Appli	187	19	95.0	356	4	US-09-328-352-7671	Sequence 7671, A
115	19	95.0	250	3	US-08-654-025-7	Sequence 7, Appli	188	19	95.0	357	4	US-09-540-236-3808	Sequence 3808, Ap
116	19	95.0	251	4	US-09-252-991A-23816	Sequence 23816, A	189	19	95.0	363	4	US-09-252-991A-30015	Sequence 30015, A
117	19	95.0	253	4	US-09-252-991A-17352	Sequence 17352, A	190	19	95.0	364	4	US-09-543-681A-7525	Sequence 7525, Ap
118	19	95.0	253	4	US-09-107-532A-5072	Sequence 5072, Ap	191	19	95.0	368	4	US-09-252-991A-24621	Sequence 24621, A
119	19	95.0	254	4	US-09-372-422A-34	Sequence 34, Appl	192	19	95.0	368	4	US-09-489-039A-9556	Sequence 9556, Ap
120	19	95.0	257	2	US-08-506-340A-3	Sequence 3, Appli	193	19	95.0	370	4	US-09-540-224-4	Sequence 4, Appli
121	19	95.0	257	4	US-09-543-681A-6811	Sequence 6811, Ap	194	19	95.0	370	4	US-09-554-595D-53	Sequence 53, Appl
122	19	95.0	259	4	US-09-543-681A-6390	Sequence 6390, Ap	195	19	95.0	370	4	US-09-808-972-4	Sequence 4, Appli
123	19	95.0	259	4	US-09-489-039A-11936	Sequence 11936, A	196	19	95.0	371	2	US-08-837-593-6	Sequence 6, Appli
124	19	95.0	261	4	US-09-252-991A-23795	Sequence 23795, A	197	19	95.0	372	4	US-09-800-729-213	Sequence 213, App
125	19	95.0	262	4	US-09-252-991A-19738	Sequence 19738, A	198	19	95.0	374	4	US-09-489-039A-8170	Sequence 8170, Ap
126	19	95.0	263	4	US-09-800-729-88	Sequence 88, Appl	199	19	95.0	388	4	US-09-489-039A-12678	Sequence 12678, A
127	19	95.0	269	1	US-08-447-554-5	Sequence 5, Appli	200	19	95.0	390	4	US-09-933-313B-8	Sequence 8, Appli
128	19	95.0	269	1	US-08-468-763-17	Sequence 17, Appl	201	19	95.0	390	4	US-08-311-731A-332	Sequence 332, App
129	19	95.0	269	1	US-08-448-160-5	Sequence 5, Appli	202	19	95.0	390	4	US-09-543-681A-6415	Sequence 6415, Ap
130	19	95.0	269	2	US-08-393-986A-17	Sequence 17, Appl	203	19	95.0	390	4	US-09-489-039A-13547	Sequence 13547, A
131	19	95.0	270	4	US-09-252-991A-29192	Sequence 29192, A	204	19	95.0	391	2	US-08-928-692-26	Sequence 26, Appl
132	19	95.0	271	4	US-09-107-532A-5071	Sequence 5071, Ap	205	19	95.0	391	4	US-09-339-972-26	Sequence 26, Appl
133	19	95.0	273	4	US-09-489-039A-9190	Sequence 9190, Ap	206	19	95.0	394	4	US-09-934-903-4	Sequence 4, Appli
134	19	95.0	277	1	US-08-400-413-1	Sequence 1, Appli	207	19	95.0	397	4	US-09-252-991A-22235	Sequence 22235, A
135	19	95.0	280	1	US-08-595-559-3	Sequence 3, Appli	208	19	95.0	398	4	US-09-242-859A-4	Sequence 4, Appli
136	19	95.0	281	4	US-09-314-701-58	Sequence 58, Appl	209	19	95.0	398	4	US-09-242-859A-8	Sequence 8, Appli
137	19	95.0	282	4	US-09-134-000C-3572	Sequence 3572, Ap	210	19	95.0	398	4	US-09-252-991A-17379	Sequence 17379, A
138	19	95.0	283	4	US-08-956-171E-5203	Sequence 5203, Ap	211	19	95.0	400	2	US-08-733-825-2	Sequence 2, Appli
139	19	95.0	288	4	US-09-252-991A-18721	Sequence 18721, A	212	19	95.0	400	3	US-09-264-097-6	Sequence 6, Appli
140	19	95.0	290	4	US-09-743-847-2	Sequence 2, Appli	213	19	95.0	403	4	US-09-540-236-2573	Sequence 2573, Ap
141	19	95.0	291	4	US-09-252-991A-25517	Sequence 25517, A	214	19	95.0	405	3	US-09-232-200-63	Sequence 63, Appl
142	19	95.0	291	4	US-09-489-039A-9578	Sequence 9578, Ap	215	19	95.0	405	4	US-09-232-197-63	Sequence 63, Appl
143	19	95.0	294	4	US-09-523-263B-20	Sequence 20, Appl	216	19	95.0	405	4	US-09-231-023A-20	Sequence 20, Appl
144	19	95.0	295	4	US-09-543-681A-5660	Sequence 5660, Ap	217	19	95.0	405	4	US-09-232-201-63	Sequence 63, Appl
145	19	95.0	297	4	US-09-489-039A-9087	Sequence 9087, Ap	218	19	95.0	405	4	US-09-540-715A-20	Sequence 20, Appl
146	19	95.0	299	4	US-09-314-701-46	Sequence 46, Appl	219	19	95.0	405	4	US-09-232-195-63	Sequence 63, Appl
147	19	95.0	299	4	US-09-393-634-35	Sequence 35, Appl	220	19	95.0	406	4	US-09-543-681A-7962	Sequence 7962, Ap
148	19	95.0	301	4	US-09-489-039A-11964	Sequence 11964, A	221	19	95.0	410	2	US-08-723-415B-10	Sequence 10, Appl
149	19	95.0	311	4	US-09-614-912-198	Sequence 198, App	222	19	95.0	410	2	US-08-723-415B-11	Sequence 11, Appl
150	19	95.0	317	4	US-09-483-077A-17	Sequence 17, Appl	223	19	95.0	410	2	US-08-428-131-2	Sequence 2, Appli
151	19	95.0	318	2	US-08-872-719-2	Sequence 2, Appli	224	19	95.0	410	3	US-09-078-596-2	Sequence 2, Appli
152	19	95.0	318	3	US-08-957-302A-12	Sequence 12, Appl	225	19	95.0	410	3	US-09-189-627A-10	Sequence 10, Appl
153	19	95.0	318	3	US-09-336-890-2	Sequence 2, Appli	226	19	95.0	410	3	US-09-189-627A-11	Sequence 11, Appl
154	19	95.0	318	3	US-09-542-403-12	Sequence 12, Appl	227	19	95.0	410	4	US-09-710-861-10	Sequence 10, Appl
155	19	95.0	318	4	US-09-668-499-2	Sequence 2, Appli	228	19	95.0	410	4	US-09-710-861-11	Sequence 11, Appl
156	19	95.0	318	4	US-09-975-594-938	Sequence 938, App	229	19	95.0	410	4	US-09-252-991A-24839	Sequence 24839, A
157	19	95.0	319	2	US-08-798-927-4	Sequence 4, Appli	230	19	95.0	411	4	US-09-540-236-3549	Sequence 3549, Ap
158	19	95.0	319	3	US-09-345-468-5	Sequence 5, Appli	231	19	95.0	411	4	US-09-252-991A-17176	Sequence 17176, A
159	19	95.0	319	4	US-09-414-453A-5	Sequence 5, Appli	232	19	95.0	414	4	US-09-134-001C-5101	Sequence 5101, Ap
160	19	95.0	321	4	US-09-440-597-4	Sequence 4, Appli	233	19	95.0	415	4	US-09-252-991A-31684	Sequence 31684, A
161	19	95.0	321	4	US-09-252-991A-18807	Sequence 18807, A	234	19	95.0	415	4	US-09-252-991A-26713	Sequence 26713, A
162	19	95.0	323	4	US-09-543-681A-17111	Sequence 17111, A	235	19	95.0	416	4	US-09-252-991A-26713	Sequence 26713, A
163	19	95.0	323	4	US-09-543-681A-6957	Sequence 6957, Ap	236	19	95.0	424	4	US-09-328-352-4187	Sequence 4187, Ap
164	19	95.0	323	4	US-09-489-039A-7408	Sequence 7408, Ap	237	19	95.0	424	4	US-09-489-039A-9628	Sequence 9628, Ap
165	19	95.0	324	4	US-09-328-352-4636	Sequence 4636, Ap	238	19	95.0	425	4	US-09-489-039A-8386	Sequence 8386, Ap
166	19	95.0	326	4	US-09-230-196-20	Sequence 20, Appl	239	19	95.0	431	4	US-09-134-001C-4357	Sequence 4357, Ap
167	19	95.0	326	4	US-09-489-039A-9512	Sequence 9512, Ap	240	19	95.0	431	4	US-09-252-991A-24878	Sequence 24878, A
168	19	95.0	327	4	US-09-252-991A-30065	Sequence 30065, A	241	19	95.0	432	1	US-08-522-166-8	Sequence 8, Appli
169	19	95.0	328	4	US-09-489-039A-9306	Sequence 9306, Ap	242	19	95.0	432	1	US-08-488-382A-8	Sequence 8, Appli
170	19	95.0	331	4	US-09-489-039A-9639	Sequence 9639, Ap	243	19	95.0	432	2	US-08-480-912-8	Sequence 8, Appli
171	19	95.0	335	1	US-08-202-054-2	Sequence 2, Appli	244	19	95.0	435	4	US-09-252-991A-30562	Sequence 30562, A
172	19	95.0	335	4	US-08-446-923-2	Sequence 2, Appli	245	19	95.0	436	4	US-09-151-984A-9	Sequence 9, Appli
173	19	95.0	335	4	US-09-482-273-118	Sequence 118, App	246	19	95.0	436	4	US-09-134-000C-3950	Sequence 3950, Ap

247	19	95.0	436	6	5405943-4	Patent No. 5405943	320	19	95.0	485	4	US-09-540-715A-18	Sequence 18, Appl
248	19	95.0	437	4	US-09-252-991A-25332	Sequence 25332, A	321	19	95.0	485	4	US-09-540-715A-19	Sequence 19, Appl
249	19	95.0	440	4	US-09-252-991A-22487	Sequence 22487, A	322	19	95.0	485	4	US-09-769-864-1	Sequence 1, Appl
250	19	95.0	444	4	US-09-252-991A-18256	Sequence 18256, A	323	19	95.0	485	4	US-09-769-864-2	Sequence 2, Appl
251	19	95.0	444	4	US-09-252-991A-19565	Sequence 19565, A	324	19	95.0	485	4	US-09-769-864-6	Sequence 6, Appl
252	19	95.0	445	4	US-09-252-991A-28986	Sequence 28986, A	325	19	95.0	485	4	US-09-769-864-7	Sequence 7, Appl
253	19	95.0	445	4	US-09-107-532A-7232	Sequence 7232, Ap	326	19	95.0	485	4	US-09-769-864-8	Sequence 8, Appl
254	19	95.0	446	4	US-09-328-352-6076	Sequence 6076, Ap	327	19	95.0	486	4	US-09-381-687-4	Sequence 4, Appl
255	19	95.0	449	4	US-09-134-000C-5444	Sequence 5444, Ap	328	19	95.0	493	3	US-08-986-441B-70	Sequence 70, Appl
256	19	95.0	450	4	US-09-543-681A-7260	Sequence 7260, Ap	329	19	95.0	493	3	US-08-993-722A-70	Sequence 70, Appl
257	19	95.0	451	4	US-09-134-000C-6421	Sequence 6421, Ap	330	19	95.0	493	3	US-08-993-170A-70	Sequence 70, Appl
258	19	95.0	455	4	US-09-252-991A-19936	Sequence 19936, A	331	19	95.0	493	3	US-08-993-775B-70	Sequence 70, Appl
259	19	95.0	456	4	US-09-252-991A-31884	Sequence 31884, A	332	19	95.0	493	3	US-08-177-349-5	Sequence 5, Appl
260	19	95.0	457	4	US-09-252-991A-28824	Sequence 28824, A	333	19	95.0	493	4	US-09-328-352-6753	Sequence 6753, Ap
261	19	95.0	467	4	US-09-543-681A-6871	Sequence 6871, Ap	334	19	95.0	493	4	US-09-427-770-70	Sequence 70, Appl
262	19	95.0	468	4	US-09-252-991A-18608	Sequence 18608, A	335	19	95.0	493	4	US-09-540-236-2869	Sequence 2869, Ap
263	19	95.0	469	4	US-09-543-681A-5423	Sequence 5423, Ap	336	19	95.0	493	4	US-09-179-966-1	Sequence 1, Appl
264	19	95.0	474	4	US-09-252-991A-27524	Sequence 27524, A	337	19	95.0	495	3	US-08-980-984-1	Sequence 1, Appl
265	19	95.0	475	4	US-09-252-991A-32806	Sequence 32806, A	338	19	95.0	495	3	US-09-252-991A-20317	Sequence 20317, A
266	19	95.0	479	4	US-09-328-352-5396	Sequence 5396, Ap	339	19	95.0	506	3	US-09-232-191-9	Sequence 9, Appl
267	19	95.0	485	2	US-08-446-803-1	Sequence 1, Appl	340	19	95.0	506	3	US-09-232-200-9	Sequence 9, Appl
268	19	95.0	485	2	US-08-446-803-2	Sequence 2, Appl	341	19	95.0	506	3	US-09-232-200-95	Sequence 9, Appl
269	19	95.0	485	2	US-08-861-837-1	Sequence 1, Appl	342	19	95.0	506	4	US-09-232-197-9	Sequence 9, Appl
270	19	95.0	485	2	US-08-861-837-2	Sequence 2, Appl	343	19	95.0	506	4	US-09-232-197-95	Sequence 9, Appl
271	19	95.0	485	2	US-08-600-908A-12	Sequence 12, Appl	344	19	95.0	506	4	US-09-232-201-9	Sequence 9, Appl
272	19	95.0	485	3	US-08-683-838A-12	Sequence 12, Appl	345	19	95.0	506	4	US-09-232-201-95	Sequence 9, Appl
273	19	95.0	485	3	US-08-600-656-1	Sequence 1, Appl	346	19	95.0	506	4	US-09-232-195-9	Sequence 9, Appl
274	19	95.0	485	3	US-08-600-656-2	Sequence 2, Appl	347	19	95.0	506	4	US-09-232-195-95	Sequence 9, Appl
275	19	95.0	485	3	US-08-600-656-7	Sequence 7, Appl	348	19	95.0	516	4	US-09-986-676A-2	Sequence 2, Appl
276	19	95.0	485	3	US-09-170-670-1	Sequence 1, Appl	349	19	95.0	522	4	US-09-252-991A-30451	Sequence 30451, A
277	19	95.0	485	3	US-09-170-670-2	Sequence 2, Appl	350	19	95.0	522	4	US-08-928-632-12	Sequence 12, Appl
278	19	95.0	485	3	US-09-170-670-6	Sequence 6, Appl	351	19	95.0	524	3	US-08-957-302A-2	Sequence 2, Appl
279	19	95.0	485	3	US-09-170-670-7	Sequence 7, Appl	352	19	95.0	524	3	US-09-542-403-2	Sequence 12, Appl
280	19	95.0	485	3	US-09-193-068-1	Sequence 1, Appl	353	19	95.0	524	3	US-09-339-972-12	Sequence 12, Appl
281	19	95.0	485	3	US-09-193-068-2	Sequence 2, Appl	354	19	95.0	532	4	US-09-252-991A-27288	Sequence 27288, A
282	19	95.0	485	3	US-09-193-068-6	Sequence 6, Appl	355	19	95.0	536	4	US-09-107-532A-6930	Sequence 6930, Ap
283	19	95.0	485	3	US-09-193-068-7	Sequence 7, Appl	356	19	95.0	536	4	US-09-489-039A-13518	Sequence 13518, A
284	19	95.0	485	3	US-09-264-097-5	Sequence 5, Appl	357	19	95.0	538	4	US-09-621-451-2	Sequence 2, Appl
285	19	95.0	485	3	US-09-193-068-8	Sequence 8, Appl	358	19	95.0	540	4	US-10-223-355-2	Sequence 2, Appl
286	19	95.0	485	3	US-09-183-412-1	Sequence 1, Appl	359	19	95.0	540	4	US-09-328-352-446	Sequence 149, App
287	19	95.0	485	3	US-09-183-412-2	Sequence 2, Appl	360	19	95.0	544	3	US-08-483-577A-149	Sequence 149, App
288	19	95.0	485	3	US-09-183-412-6	Sequence 6, Appl	361	19	95.0	547	3	US-08-897-438-149	Sequence 149, App
289	19	95.0	485	3	US-09-183-412-7	Sequence 7, Appl	362	19	95.0	547	3	US-08-649-518-149	Sequence 149, App
290	19	95.0	485	3	US-09-183-412-8	Sequence 8, Appl	363	19	95.0	547	4	US-09-396-478A-2	Sequence 2, Appl
291	19	95.0	485	3	US-09-264-097-5	Sequence 5, Appl	364	19	95.0	550	4	US-09-489-039A-13735	Sequence 13735, A
292	19	95.0	485	3	US-09-264-097-7	Sequence 7, Appl	365	19	95.0	552	4	US-09-134-001C-5569	Sequence 5569, Ap
293	19	95.0	485	3	US-09-354-191A-1	Sequence 1, Appl	366	19	95.0	557	4	US-09-489-039A-13795	Sequence 13795, A
294	19	95.0	485	3	US-09-354-191A-2	Sequence 2, Appl	367	19	95.0	557	4	US-09-489-039A-13795	Sequence 4354, Ap
295	19	95.0	485	3	US-09-354-191A-7	Sequence 7, Appl	368	19	95.0	557	4	US-09-252-991A-27678	Sequence 27678, A
296	19	95.0	485	4	US-09-291-023A-13	Sequence 13, Appl	369	19	95.0	558	4	US-09-513-783A-142	Sequence 142, App
297	19	95.0	485	4	US-09-291-023A-18	Sequence 18, Appl	370	19	95.0	566	4	US-08-504-459-4	Sequence 4, Appl
298	19	95.0	485	4	US-09-291-023A-19	Sequence 19, Appl	371	19	95.0	567	2	US-08-504-459-6	Sequence 6, Appl
299	19	95.0	485	4	US-09-290-734-1	Sequence 1, Appl	372	19	95.0	567	2	US-08-803-326A-6	Sequence 6, Appl
300	19	95.0	485	4	US-09-290-734-2	Sequence 2, Appl	373	19	95.0	571	3	US-09-689-913A-6	Sequence 6, Appl
301	19	95.0	485	4	US-09-290-734-6	Sequence 6, Appl	374	19	95.0	571	4	US-09-689-913A-6	Sequence 6, Appl
302	19	95.0	485	4	US-09-290-734-7	Sequence 7, Appl	375	19	95.0	571	4	US-09-689-916A-6	Sequence 6, Appl
303	19	95.0	485	4	US-09-290-734-8	Sequence 8, Appl	376	19	95.0	571	4	US-09-252-991A-18744	Sequence 18744, A
304	19	95.0	485	4	US-09-290-734-24	Sequence 24, Appl	377	19	95.0	573	4	US-09-328-352-4675	Sequence 332, App
305	19	95.0	485	4	US-09-290-734-26	Sequence 26, Appl	378	19	95.0	580	4	US-09-198-452A-332	Sequence 3, Appl
306	19	95.0	485	4	US-09-636-252A-12	Sequence 12, Appl	379	19	95.0	580	4	US-08-635-552A-3	Sequence 4, Appl
307	19	95.0	485	4	US-09-417-359A-5	Sequence 5, Appl	380	19	95.0	587	4	US-08-675-439A-4	Sequence 4, Appl
308	19	95.0	485	4	US-09-381-687-1	Sequence 1, Appl	381	19	95.0	587	4	US-08-812-008-4	Sequence 4, Appl
309	19	95.0	485	4	US-09-381-687-2	Sequence 2, Appl	382	19	95.0	587	4	US-09-643-657-14	Sequence 14, Appl
310	19	95.0	485	4	US-09-381-687-3	Sequence 3, Appl	383	19	95.0	589	4	US-08-736-770-5	Sequence 5, Appl
311	19	95.0	485	4	US-09-381-687-5	Sequence 5, Appl	384	19	95.0	591	2	US-09-643-657-15	Sequence 15, Appl
312	19	95.0	485	4	US-09-545-586-1	Sequence 1, Appl	385	19	95.0	591	4	US-09-545-586-2	Sequence 2, Appl
313	19	95.0	485	4	US-09-545-586-2	Sequence 2, Appl	386	19	95.0	591	4	US-09-545-586-6	Sequence 6, Appl
314	19	95.0	485	4	US-09-545-586-7	Sequence 7, Appl	387	19	95.0	597	4	US-09-252-991A-23152	Sequence 23152, A
315	19	95.0	485	4	US-09-545-586-8	Sequence 8, Appl	388	19	95.0	597	4	US-08-472-534-5	Sequence 5, Appl
316	19	95.0	485	4	US-09-545-586-24	Sequence 24, Appl	389	19	95.0	607	2	US-08-637-670-40	Sequence 40, Appl
317	19	95.0	485	4	US-09-545-586-26	Sequence 26, Appl	390	19	95.0	608	2	US-09-041-991A-6	Sequence 6, Appl
318	19	95.0	485	4	US-09-540-715A-13	Sequence 13, Appl	391	19	95.0	620	4		
319	19	95.0	485	4			392	19	95.0	623	3		

393	19	95.0	623	4	US-08-608-533A-6	Sequence 6, Appl	466	19	95.0	651	3	US-08-993-170A-56	Sequence 56, Appl
394	19	95.0	625	4	US-08-611-322A-48	Sequence 48, Appl	467	19	95.0	651	3	US-08-993-170A-58	Sequence 58, Appl
395	19	95.0	626	4	US-08-232-891A-20268	Sequence 20268, A	468	19	95.0	651	3	US-08-993-775B-52	Sequence 52, Appl
396	19	95.0	631	1	US-08-487-890A-111	Sequence 111, Appl	469	19	95.0	651	3	US-08-993-775B-56	Sequence 56, Appl
397	19	95.0	631	2	US-08-478-435-111	Sequence 111, Appl	470	19	95.0	651	3	US-08-993-775B-58	Sequence 58, Appl
398	19	95.0	631	2	US-08-337-483-111	Sequence 111, Appl	471	19	95.0	651	4	US-09-427-770-52	Sequence 52, Appl
399	19	95.0	631	2	US-08-478-373-111	Sequence 111, Appl	472	19	95.0	651	4	US-09-427-770-56	Sequence 56, Appl
400	19	95.0	631	3	US-08-474-671-111	Sequence 111, Appl	473	19	95.0	651	4	US-09-427-769-52	Sequence 52, Appl
401	19	95.0	631	3	US-08-483-577A-111	Sequence 111, Appl	474	19	95.0	651	4	US-09-427-769-56	Sequence 56, Appl
402	19	95.0	631	3	US-08-897-438-111	Sequence 111, Appl	475	19	95.0	651	4	US-09-427-769-58	Sequence 58, Appl
403	19	95.0	631	4	US-08-637-654-111	Sequence 111, Appl	476	19	95.0	651	4	US-09-427-769-58	Sequence 58, Appl
404	19	95.0	631	4	US-08-649-518-111	Sequence 111, Appl	477	19	95.0	652	3	US-08-996-441B-2	Sequence 2, Appl
405	19	95.0	632	3	US-09-232-200-34	Sequence 34, Appl	478	19	95.0	652	3	US-08-996-441B-4	Sequence 4, Appl
406	19	95.0	632	3	US-09-232-200-35	Sequence 35, Appl	479	19	95.0	652	3	US-08-996-441B-6	Sequence 6, Appl
407	19	95.0	632	3	US-09-232-200-39	Sequence 39, Appl	480	19	95.0	652	3	US-08-996-441B-8	Sequence 8, Appl
408	19	95.0	632	4	US-09-232-197-34	Sequence 34, Appl	481	19	95.0	652	3	US-08-996-441B-10	Sequence 10, Appl
409	19	95.0	632	4	US-09-232-197-35	Sequence 35, Appl	482	19	95.0	652	3	US-08-996-441B-12	Sequence 12, Appl
410	19	95.0	632	4	US-09-232-197-39	Sequence 39, Appl	483	19	95.0	652	3	US-08-996-441B-14	Sequence 14, Appl
411	19	95.0	632	4	US-09-232-201-34	Sequence 34, Appl	484	19	95.0	652	3	US-08-996-441B-16	Sequence 16, Appl
412	19	95.0	632	4	US-09-232-201-35	Sequence 35, Appl	485	19	95.0	652	3	US-08-996-441B-18	Sequence 18, Appl
413	19	95.0	632	4	US-09-232-201-39	Sequence 39, Appl	486	19	95.0	652	3	US-08-996-441B-20	Sequence 20, Appl
414	19	95.0	632	4	US-09-232-195-34	Sequence 34, Appl	487	19	95.0	652	3	US-08-996-441B-22	Sequence 22, Appl
415	19	95.0	632	4	US-09-232-195-35	Sequence 35, Appl	488	19	95.0	652	3	US-08-996-441B-24	Sequence 24, Appl
416	19	95.0	632	4	US-09-232-195-39	Sequence 39, Appl	489	19	95.0	652	3	US-08-996-441B-26	Sequence 26, Appl
417	19	95.0	635	4	US-08-107-322A-7135	Sequence 7135, Ap	490	19	95.0	652	3	US-08-996-441B-28	Sequence 28, Appl
418	19	95.0	638	4	US-08-489-039A-8996	Sequence 8996, Ap	491	19	95.0	652	3	US-08-996-441B-30	Sequence 30, Appl
419	19	95.0	639	4	US-08-328-352-4979	Sequence 4979, Ap	492	19	95.0	652	3	US-08-996-441B-32	Sequence 32, Appl
420	19	95.0	640	4	US-08-252-991A-23007	Sequence 23007, A	493	19	95.0	652	3	US-08-996-441B-34	Sequence 34, Appl
421	19	95.0	641	4	US-09-543-681A-6258	Sequence 6258, Ap	494	19	95.0	652	3	US-08-996-441B-36	Sequence 36, Appl
422	19	95.0	643	3	US-09-232-200-27	Sequence 27, Appl	495	19	95.0	652	3	US-08-996-441B-38	Sequence 38, Appl
423	19	95.0	643	3	US-09-232-200-41	Sequence 41, Appl	496	19	95.0	652	3	US-08-996-441B-40	Sequence 40, Appl
424	19	95.0	643	3	US-09-232-200-42	Sequence 42, Appl	497	19	95.0	652	3	US-08-996-441B-42	Sequence 42, Appl
425	19	95.0	643	3	US-09-232-200-45	Sequence 45, Appl	498	19	95.0	652	3	US-08-996-441B-44	Sequence 44, Appl
426	19	95.0	643	3	US-09-232-200-53	Sequence 53, Appl	499	19	95.0	652	3	US-08-996-441B-46	Sequence 46, Appl
427	19	95.0	643	3	US-09-232-200-71	Sequence 71, Appl	500	19	95.0	652	3	US-08-996-441B-48	Sequence 48, Appl
428	19	95.0	643	4	US-09-232-197-27	Sequence 27, Appl	501	19	95.0	652	3	US-08-996-441B-50	Sequence 50, Appl
429	19	95.0	643	4	US-09-232-197-42	Sequence 42, Appl	502	19	95.0	652	3	US-08-996-441B-54	Sequence 54, Appl
430	19	95.0	643	4	US-09-232-197-45	Sequence 45, Appl	503	19	95.0	652	3	US-08-996-441B-60	Sequence 60, Appl
431	19	95.0	643	4	US-09-232-197-53	Sequence 53, Appl	504	19	95.0	652	3	US-08-996-441B-62	Sequence 62, Appl
432	19	95.0	643	4	US-09-232-197-71	Sequence 71, Appl	505	19	95.0	652	3	US-08-996-441B-64	Sequence 64, Appl
433	19	95.0	643	4	US-09-232-201-27	Sequence 27, Appl	506	19	95.0	652	3	US-08-996-441B-66	Sequence 66, Appl
434	19	95.0	643	4	US-09-232-201-41	Sequence 41, Appl	507	19	95.0	652	3	US-08-996-441B-68	Sequence 68, Appl
435	19	95.0	643	4	US-09-232-201-42	Sequence 42, Appl	508	19	95.0	652	3	US-08-996-441B-98	Sequence 98, Appl
436	19	95.0	643	4	US-09-232-201-45	Sequence 45, Appl	509	19	95.0	652	3	US-08-996-441B-108	Sequence 108, Appl
437	19	95.0	643	4	US-09-232-201-53	Sequence 53, Appl	510	19	95.0	652	3	US-08-996-441B-110	Sequence 110, Appl
438	19	95.0	643	4	US-08-232-201-71	Sequence 71, Appl	511	19	95.0	652	3	US-08-996-441B-111	Sequence 111, Appl
439	19	95.0	643	4	US-09-543-661A-6345	Sequence 6345, Ap	512	19	95.0	652	3	US-08-993-722A-2	Sequence 2, Appl
440	19	95.0	643	4	US-09-232-195-27	Sequence 27, Appl	513	19	95.0	652	3	US-08-993-722A-4	Sequence 4, Appl
441	19	95.0	643	4	US-09-232-195-41	Sequence 41, Appl	514	19	95.0	652	3	US-08-993-722A-6	Sequence 6, Appl
442	19	95.0	643	4	US-09-232-195-42	Sequence 42, Appl	515	19	95.0	652	3	US-08-993-722A-8	Sequence 8, Appl
443	19	95.0	643	4	US-09-232-195-45	Sequence 45, Appl	516	19	95.0	652	3	US-08-993-722A-10	Sequence 10, Appl
444	19	95.0	643	4	US-09-232-195-53	Sequence 53, Appl	517	19	95.0	652	3	US-08-993-722A-12	Sequence 12, Appl
445	19	95.0	643	4	US-08-232-195-71	Sequence 71, Appl	518	19	95.0	652	3	US-08-993-722A-14	Sequence 14, Appl
446	19	95.0	644	1	US-08-487-890A-6	Sequence 6, Appl	519	19	95.0	652	3	US-08-993-722A-16	Sequence 16, Appl
447	19	95.0	644	2	US-08-478-435-6	Sequence 6, Appl	520	19	95.0	652	3	US-08-993-722A-18	Sequence 18, Appl
448	19	95.0	644	2	US-08-337-483-6	Sequence 6, Appl	521	19	95.0	652	3	US-08-993-722A-20	Sequence 20, Appl
449	19	95.0	644	2	US-08-478-373-6	Sequence 6, Appl	522	19	95.0	652	3	US-08-993-722A-22	Sequence 22, Appl
450	19	95.0	644	3	US-08-474-671-6	Sequence 6, Appl	523	19	95.0	652	3	US-08-993-722A-24	Sequence 24, Appl
451	19	95.0	644	3	US-08-483-577A-6	Sequence 6, Appl	524	19	95.0	652	3	US-08-993-722A-26	Sequence 26, Appl
452	19	95.0	644	3	US-08-897-438-6	Sequence 6, Appl	525	19	95.0	652	3	US-08-993-722A-28	Sequence 28, Appl
453	19	95.0	644	4	US-08-637-654-6	Sequence 6, Appl	526	19	95.0	652	3	US-08-993-722A-30	Sequence 30, Appl
454	19	95.0	644	4	US-08-649-518-6	Sequence 6, Appl	527	19	95.0	652	3	US-08-993-722A-32	Sequence 32, Appl
455	19	95.0	644	4	US-08-897-438-6	Sequence 6, Appl	528	19	95.0	652	3	US-08-993-722A-34	Sequence 34, Appl
456	19	95.0	647	3	US-08-483-577A-148	Sequence 148, App	529	19	95.0	652	3	US-08-993-722A-36	Sequence 36, Appl
457	19	95.0	647	3	US-08-897-438-148	Sequence 148, App	530	19	95.0	652	3	US-08-993-722A-38	Sequence 38, Appl
458	19	95.0	647	4	US-08-649-518-148	Sequence 148, App	531	19	95.0	652	3	US-08-993-722A-40	Sequence 40, Appl
459	19	95.0	651	3	US-08-996-441B-52	Sequence 52, Appl	532	19	95.0	652	3	US-08-993-722A-42	Sequence 42, Appl
460	19	95.0	651	3	US-08-996-441B-56	Sequence 56, Appl	533	19	95.0	652	3	US-08-993-722A-44	Sequence 44, Appl
461	19	95.0	651	3	US-08-996-441B-58	Sequence 58, Appl	534	19	95.0	652	3	US-08-993-722A-46	Sequence 46, Appl
462	19	95.0	651	3	US-08-993-722A-52	Sequence 52, Appl	535	19	95.0	652	3	US-08-993-722A-48	Sequence 48, Appl
463	19	95.0	651	3	US-08-993-722A-56	Sequence 56, Appl	536	19	95.0	652	3	US-08-993-722A-50	Sequence 50, Appl
464	19	95.0	651	3	US-08-993-722A-58	Sequence 58, Appl	537	19	95.0	652	3	US-08-993-722A-54	Sequence 54, Appl
465	19	95.0	651	3	US-08-993-170A-52	Sequence 52, Appl	538	19	95.0	652	3	US-08-993-722A-60	Sequence 60, Appl

539	19	95.0	652	3	US-08-993-722A-62	Sequence 62, Appl	612	19	95.0	652	3	US-08-993-775B-68	Sequence 68, Appl
540	19	95.0	652	3	US-08-993-722A-64	Sequence 64, Appl	613	19	95.0	652	3	US-08-993-775B-98	Sequence 98, Appl
541	19	95.0	652	3	US-08-993-722A-66	Sequence 66, Appl	614	19	95.0	652	3	US-08-993-775B-108	Sequence 108, Appl
542	19	95.0	652	3	US-08-993-722A-68	Sequence 68, Appl	615	19	95.0	652	3	US-08-993-775B-110	Sequence 110, Appl
543	19	95.0	652	3	US-08-993-722A-98	Sequence 98, Appl	616	19	95.0	652	3	US-08-993-775B-111	Sequence 111, Appl
544	19	95.0	652	3	US-08-993-722A-108	Sequence 108, Appl	617	19	95.0	652	4	US-09-377-466B-2	Sequence 2, Appl
545	19	95.0	652	3	US-08-993-722A-110	Sequence 110, Appl	618	19	95.0	652	4	US-09-377-466B-4	Sequence 4, Appl
546	19	95.0	652	3	US-08-993-722A-111	Sequence 111, Appl	619	19	95.0	652	4	US-09-377-466B-6	Sequence 6, Appl
547	19	95.0	652	3	US-08-993-170A-2	Sequence 2, Appl	620	19	95.0	652	4	US-09-427-770-4	Sequence 4, Appl
548	19	95.0	652	3	US-08-993-170A-6	Sequence 6, Appl	621	19	95.0	652	4	US-09-427-770-6	Sequence 6, Appl
549	19	95.0	652	3	US-08-993-170A-8	Sequence 8, Appl	622	19	95.0	652	4	US-09-427-770-8	Sequence 8, Appl
550	19	95.0	652	3	US-08-993-170A-10	Sequence 10, Appl	623	19	95.0	652	4	US-09-427-770-10	Sequence 10, Appl
551	19	95.0	652	3	US-08-993-170A-12	Sequence 12, Appl	624	19	95.0	652	4	US-09-427-770-12	Sequence 12, Appl
552	19	95.0	652	3	US-08-993-170A-14	Sequence 14, Appl	625	19	95.0	652	4	US-09-427-770-14	Sequence 14, Appl
553	19	95.0	652	3	US-08-993-170A-16	Sequence 16, Appl	626	19	95.0	652	4	US-09-427-770-16	Sequence 16, Appl
554	19	95.0	652	3	US-08-993-170A-18	Sequence 18, Appl	627	19	95.0	652	4	US-09-427-770-18	Sequence 18, Appl
555	19	95.0	652	3	US-08-993-170A-20	Sequence 20, Appl	628	19	95.0	652	4	US-09-427-770-20	Sequence 20, Appl
556	19	95.0	652	3	US-08-993-170A-22	Sequence 22, Appl	629	19	95.0	652	4	US-09-427-770-22	Sequence 22, Appl
557	19	95.0	652	3	US-08-993-170A-24	Sequence 24, Appl	630	19	95.0	652	4	US-09-427-770-24	Sequence 24, Appl
558	19	95.0	652	3	US-08-993-170A-26	Sequence 26, Appl	631	19	95.0	652	4	US-09-427-770-26	Sequence 26, Appl
559	19	95.0	652	3	US-08-993-170A-28	Sequence 28, Appl	632	19	95.0	652	4	US-09-427-770-28	Sequence 28, Appl
560	19	95.0	652	3	US-08-993-170A-30	Sequence 30, Appl	633	19	95.0	652	4	US-09-427-770-30	Sequence 30, Appl
561	19	95.0	652	3	US-08-993-170A-32	Sequence 32, Appl	634	19	95.0	652	4	US-09-427-770-32	Sequence 32, Appl
562	19	95.0	652	3	US-08-993-170A-34	Sequence 34, Appl	635	19	95.0	652	4	US-09-427-770-34	Sequence 34, Appl
563	19	95.0	652	3	US-08-993-170A-36	Sequence 36, Appl	636	19	95.0	652	4	US-09-427-770-36	Sequence 36, Appl
564	19	95.0	652	3	US-08-993-170A-38	Sequence 38, Appl	637	19	95.0	652	4	US-09-427-770-38	Sequence 38, Appl
565	19	95.0	652	3	US-08-993-170A-40	Sequence 40, Appl	638	19	95.0	652	4	US-09-427-770-40	Sequence 40, Appl
566	19	95.0	652	3	US-08-993-170A-42	Sequence 42, Appl	639	19	95.0	652	4	US-09-427-770-42	Sequence 42, Appl
567	19	95.0	652	3	US-08-993-170A-								

685	19	95.0	652	4	US-09-427-769-68	Sequence 58, Appl	758	19	95.0	736	4	US-09-252-991A-27827	Sequence 27827, A
686	19	95.0	652	4	US-09-427-769-98	Sequence 98, Appl	759	19	95.0	748	4	US-09-252-991A-28449	Sequence 28449, A
687	19	95.0	652	4	US-09-427-769-108	Sequence 108, Appl	760	19	95.0	749	4	US-09-562-737-100	Sequence 100, Appl
688	19	95.0	652	4	US-09-427-769-110	Sequence 110, Appl	761	19	95.0	749	4	US-09-252-991A-16877	Sequence 16877, A
689	19	95.0	652	4	US-09-427-769-111	Sequence 111, Appl	762	19	95.0	753	4	US-09-328-352-5412	Sequence 5412, Ap
690	19	95.0	652	5	PCT-US92-00040-2	Sequence 2, Appl	763	19	95.0	771	4	US-09-252-991A-17243	Sequence 17243, A
691	19	95.0	652	5	5187091-2	Patent No. 5187091	764	19	95.0	788	2	US-08-907-166-6	Sequence 6, Appl
692	19	95.0	652	3	US-08-986-441B-100	Sequence 100, Appl	765	19	95.0	788	4	US-09-391-340-6	Sequence 16, Appl
693	19	95.0	653	3	US-08-993-722A-100	Sequence 100, Appl	766	19	95.0	789	4	US-09-390-234-16	Sequence 16, Appl
694	19	95.0	653	3	US-08-993-170A-100	Sequence 100, Appl	767	19	95.0	789	4	US-09-603-311-16	Sequence 6065, Ap
695	19	95.0	653	3	US-08-993-775B-100	Sequence 100, Appl	768	19	95.0	791	4	US-09-107-532A-065	Sequence 4, Appl
696	19	95.0	653	4	US-09-377-466B-8	Sequence 8, Appl	769	19	95.0	812	4	US-09-513-783A-4	Sequence 6, Appl
697	19	95.0	653	4	US-09-377-466B-10	Sequence 10, Appl	770	19	95.0	812	4	US-09-513-783A-6	Sequence 10, Appl
698	19	95.0	653	4	US-09-377-466B-12	Sequence 12, Appl	771	19	95.0	816	4	US-09-489-038A-10520	Sequence 15, Appl
699	19	95.0	653	4	US-09-377-466B-14	Sequence 14, Appl	772	19	95.0	819	4	US-09-651-656-15	Sequence 20, Appl
700	19	95.0	653	4	US-09-377-466B-16	Sequence 16, Appl	773	19	95.0	819	4	US-09-650-855-15	Sequence 24, Appl
701	19	95.0	653	4	US-09-377-466B-18	Sequence 18, Appl	774	19	95.0	821	3	US-09-422-869-24	Sequence 20, Appl
702	19	95.0	653	4	US-09-377-466B-20	Sequence 20, Appl	775	19	95.0	844	3	US-09-029-267-20	Sequence 20, Appl
703	19	95.0	653	4	US-09-377-466B-22	Sequence 22, Appl	776	19	95.0	846	3	US-09-328-352-6648	Sequence 6648, Ap
704	19	95.0	653	4	US-09-377-466B-37	Sequence 37, Appl	777	19	95.0	848	4	US-09-489-039A-13773	Sequence 13773, A
705	19	95.0	653	4	US-09-377-466B-39	Sequence 39, Appl	778	19	95.0	849	4	US-09-252-991A-17953	Sequence 17953, A
706	19	95.0	653	4	US-09-377-466B-100	Sequence 100, Appl	779	19	95.0	873	4	US-09-252-991A-30504	Sequence 30504, A
707	19	95.0	653	4	US-09-427-769-100	Sequence 100, Appl	780	19	95.0	925	4	US-09-252-991A-27057	Sequence 27057, A
708	19	95.0	653	4	US-09-427-769-100	Sequence 100, Appl	781	19	95.0	927	4	US-09-134-001C-4831	Sequence 4831, Ap
709	19	95.0	654	1	US-08-487-890A-12	Sequence 12, Appl	782	19	95.0	934	4	US-09-328-352-4401	Sequence 4401, Ap
710	19	95.0	654	2	US-08-478-435-12	Sequence 12, Appl	783	19	95.0	944	3	US-08-335-844A-24	Sequence 24, Appl
711	19	95.0	654	2	US-08-337-483-12	Sequence 12, Appl	784	19	95.0	972	4	US-09-129-366-24	Sequence 24, Appl
712	19	95.0	654	2	US-08-478-373-12	Sequence 12, Appl	785	19	95.0	972	4	US-09-543-681A-5755	Sequence 5755, Ap
713	19	95.0	654	3	US-08-474-671-12	Sequence 12, Appl	786	19	95.0	982	4	US-09-023-905A-7	Sequence 7, Appl
714	19	95.0	654	3	US-08-483-577A-12	Sequence 12, Appl	787	19	95.0	1074	3	US-09-252-991A-22792	Sequence 22792, A
715	19	95.0	654	3	US-08-897-438-12	Sequence 12, Appl	788	19	95.0	1093	4	US-09-894-998A-50	Sequence 50, Appl
716	19	95.0	654	4	US-08-637-654-12	Sequence 12, Appl	789	19	95.0	1037	4	US-09-693-146-4	Sequence 4, Appl
717	19	95.0	654	4	US-08-649-518-12	Sequence 12, Appl	790	19	95.0	1037	4	US-09-894-998A-54	Sequence 54, Appl
718	19	95.0	657	3	US-08-613-009A-19	Sequence 19, Appl	791	19	95.0	1036	4	US-09-513-783A-32	Sequence 32, Appl
719	19	95.0	657	4	US-08-778-570B-25	Sequence 25, Appl	792	19	95.0	1061	4	US-09-252-991A-23691	Sequence 23691, A
720	19	95.0	657	4	US-08-093-584-25	Sequence 25, Appl	793	19	95.0	1074	3	US-08-470-058-2	Sequence 2, Appl
721	19	95.0	657	4	US-09-252-991A-27358	Sequence 27358, A	794	19	95.0	1074	3	US-09-037-188-2	Sequence 2, Appl
722	19	95.0	660	1	US-08-487-890A-8	Sequence 8, Appl	795	19	95.0	1113	4	US-09-285-310-2	Sequence 2, Appl
723	19	95.0	660	1	US-08-487-890A-10	Sequence 10, Appl	796	19	95.0	1113	4	US-09-894-998A-51	Sequence 51, Appl
724	19	95.0	660	2	US-08-478-435-8	Sequence 8, Appl	797	19	95.0	1114	4	US-08-975-413A-12	Sequence 12, Appl
725	19	95.0	660	2	US-08-478-435-10	Sequence 10, Appl	798	19	95.0	1125	4	US-09-513-783A-182	Sequence 182, App
726	19	95.0	660	2	US-08-337-483-10	Sequence 10, Appl	799	19	95.0	1128	4	US-09-252-991A-19672	Sequence 19672, A
727	19	95.0	660	2	US-08-478-373-8	Sequence 8, Appl	800	19	95.0	1155	4	US-09-543-681A-6286	Sequence 6286, Ap
728	19	95.0	660	2	US-08-478-373-10	Sequence 10, Appl	801	19	95.0	1410	2	US-08-470-058-4	Sequence 4, Appl
729	19	95.0	660	2	US-08-478-373-12	Sequence 12, Appl	802	19	95.0	1410	3	US-09-037-188-4	Sequence 4, Appl
730	19	95.0	660	3	US-08-474-671-8	Sequence 8, Appl	803	19	95.0	1410	3	US-09-385-310-4	Sequence 10, Appl
731	19	95.0	660	3	US-08-474-671-10	Sequence 10, Appl	804	19	95.0	1477	3	US-08-492-459-10	Sequence 10, Appl
732	19	95.0	660	3	US-08-483-577A-8	Sequence 8, Appl	805	19	95.0	1477	3	US-08-423-752-10	Sequence 3, Appl
733	19	95.0	660	3	US-08-483-577A-10	Sequence 10, Appl	806	19	95.0	1477	3	US-08-945-994-3	Sequence 24, Appl
734	19	95.0	660	3	US-08-897-438-8	Sequence 8, Appl	807	19	95.0	1477	3	US-08-716-873-24	Sequence 24, Appl
735	19	95.0	660	3	US-08-897-438-10	Sequence 10, Appl	808	19	95.0	1477	3	US-09-368-431-24	Sequence 10, Appl
736	19	95.0	660	4	US-08-637-654-8	Sequence 8, Appl	809	19	95.0	1477	4	US-09-414-006-10	Sequence 10, Appl
737	19	95.0	660	4	US-08-637-654-10	Sequence 10, Appl	810	19	95.0	1477	4	US-09-447-223-10	Sequence 10, Appl
738	19	95.0	660	4	US-08-649-518-8	Sequence 8, Appl	811	19	95.0	1490	4	US-09-252-991A-28442	Sequence 28442, A
739	19	95.0	660	4	US-08-649-518-10	Sequence 10, Appl	812	19	95.0	1610	4	US-09-513-783A-22	Sequence 22, Appl
740	19	95.0	678	1	US-08-282-141-2	Sequence 2, Appl	813	19	95.0	2165	1	US-08-514-975B-2	Sequence 2, Appl
741	19	95.0	678	1	US-08-435-434-2	Sequence 2, Appl	814	19	95.0	2165	5	PCT-US95-12507-2	Sequence 2, Appl
742	19	95.0	678	1	US-08-435-436-2	Sequence 2, Appl	815	19	95.0	2227	3	US-08-475-886-2	Sequence 2, Appl
743	19	95.0	678	2	US-08-438-862-2	Sequence 2, Appl	816	19	95.0	2227	3	US-08-475-886-4	Sequence 6, Appl
744	19	95.0	678	3	US-08-438-864-2	Sequence 2, Appl	817	19	95.0	2227	3	US-08-475-886-6	Sequence 2, Appl
745	19	95.0	678	3	US-08-438-862-2	Sequence 2, Appl	818	19	95.0	2227	3	US-08-397-232-4	Sequence 4, Appl
746	19	95.0	678	3	US-08-628-747-2	Sequence 2, Appl	819	19	95.0	2227	3	US-08-397-232-2	Sequence 2, Appl
747	19	95.0	678	3	US-08-402-253-2	Sequence 2, Appl	820	19	95.0	2227	3	US-09-171-387-2	Sequence 2, Appl
748	19	95.0	678	3	US-08-443-866B-2	Sequence 2, Appl	821	19	95.0	2227	4	US-09-653-499-2	Sequence 4, Appl
749	19	95.0	680	4	US-09-252-991A-26639	Sequence 26639, A	822	19	95.0	2227	4	US-09-653-499-6	Sequence 6, Appl
750	19	95.0	682	4	US-09-252-991A-24919	Sequence 24919, A	823	19	95.0	2227	4	US-10-104-966-12	Sequence 12, Appl
751	19	95.0	685	4	US-09-252-991A-26723	Sequence 26723, A	824	19	95.0	2227	4	US-10-135-988-2	Sequence 2, Appl
752	19	95.0	696	6	5262177-4	Patent No. 5262177	825	19	95.0	2227	4	US-10-135-988-4	Sequence 4, Appl
753	19	95.0	698	2	US-08-175-150A-2	Sequence 2, Appl	826	19	95.0	2227	4	US-10-135-988-6	Sequence 6, Appl
754	19	95.0	703	4	US-09-252-991A-23905	Sequence 23905, A	827	19	95.0	2227	4	US-08-015-973-1	Sequence 1, Appl
755	19	95.0	718	4	US-09-252-991A-33109	Sequence 33109, A	828	19	95.0	2308	1	US-08-448-164-1	Sequence 1, Appl
756	19	95.0	720	4	US-09-252-991A-18618	Sequence 18618, A	829	19	95.0	2308	1	US-08-448-164-1	Sequence 1, Appl
757	19	95.0	720	4	US-09-134-000C-5631	Sequence 5631, Ap	830	19	95.0	2308	3	US-08-081-929-2	Sequence 2, Appl

831	19	95.0	2314	4	US-09-816-703A-2	Sequence 2, Appli	904	18	90.0	149	1	US-08-100-874-2	Sequence 2, Appli
832	19	95.0	2629	2	US-08-751-189-4	Sequence 4, Appli	905	18	90.0	149	3	US-08-963-409-4	Sequence 4, Appli
833	19	95.0	2629	2	US-09-060-836-4	Sequence 4, Appli	906	18	90.0	149	3	US-08-641-873-20	Sequence 20, Appli
834	19	95.0	2629	3	US-09-184-445-4	Sequence 4, Appli	907	18	90.0	152	3	US-08-963-409-5	Sequence 5, Appli
835	19	95.0	5588	3	US-09-036-987A-6	Sequence 6, Appli	908	18	90.0	152	4	US-09-252-991A-18700	Sequence 18700, A
836	19	95.0	5588	3	US-09-370-707-6	Sequence 6, Appli	909	18	90.0	152	4	US-09-252-991A-31418	Sequence 31418, A
837	19	95.0	5588	3	US-09-603-207-6	Sequence 6, Appli	910	18	90.0	157	4	US-09-134-001C-2991	Sequence 2991, Ap
838	19	95.0	8991	4	US-08-714-741-32	Sequence 32, Appli	911	18	90.0	158	1	US-08-611-107-4	Sequence 4, Appli
839	18	90.0	9	4	US-09-372-732-6	Sequence 6, Appli	912	18	90.0	158	2	US-08-422-560A-4	Sequence 4, Appli
840	18	90.0	9	4	US-10-126-752-6	Sequence 6, Appli	913	18	90.0	158	3	US-08-468-793-4	Sequence 4, Appli
841	18	90.0	10	2	US-08-482-228-120	Sequence 120, App	914	18	90.0	160	2	US-08-602-941-1	Sequence 1, Appli
842	18	90.0	10	2	US-08-482-228-120	Sequence 120, App	915	18	90.0	160	2	US-08-961-264-1	Sequence 1, Appli
843	18	90.0	11	2	US-08-934-222-99	Sequence 99, Appli	916	18	90.0	160	4	US-09-442-099A-1	Sequence 1, Appli
844	18	90.0	11	2	US-08-934-222-99	Sequence 99, Appli	917	18	90.0	160	4	US-09-612-342-1	Sequence 1, Appli
845	18	90.0	11	2	US-08-307-621-99	Sequence 99, Appli	918	18	90.0	160	4	US-09-612-401-1	Sequence 1, Appli
846	18	90.0	11	2	US-08-332-818-99	Sequence 99, Appli	919	18	90.0	160	4	US-09-734-401-1	Sequence 1, Appli
847	18	90.0	11	3	US-09-231-797-99	Sequence 99, Appli	920	18	90.0	162	3	US-09-108-020-50	Sequence 50, Appli
848	18	90.0	11	3	US-08-934-224-99	Sequence 99, Appli	921	18	90.0	172	4	US-09-107-532A-5847	Sequence 5847, Ap
849	18	90.0	11	3	US-08-934-224-99	Sequence 99, Appli	922	18	90.0	179	4	US-09-634-238-296	Sequence 296, App
850	18	90.0	11	3	US-08-934-223-99	Sequence 99, Appli	923	18	90.0	184	4	US-09-489-039A-11491	Sequence 11491, A
851	18	90.0	11	3	US-09-413-492-99	Sequence 99, Appli	924	18	90.0	185	1	US-07-741-940-6	Sequence 6, Appli
852	18	90.0	19	4	US-08-835-281-10	Sequence 10, Appli	925	18	90.0	185	1	US-08-289-548A-6	Sequence 6, Appli
853	18	90.0	20	1	US-08-658-130-1	Sequence 1, Appli	926	18	90.0	185	1	US-08-452-654-6	Sequence 6, Appli
854	18	90.0	20	2	US-08-799-825-22	Sequence 22, Appli	927	18	90.0	185	1	US-08-452-655B-6	Sequence 6, Appli
855	18	90.0	20	4	US-09-092-000-5	Sequence 5, Appli	928	18	90.0	185	2	US-08-865-336-3	Sequence 3, Appli
856	18	90.0	20	4	US-09-280-068-22	Sequence 22, Appli	929	18	90.0	185	2	US-08-865-336-4	Sequence 4, Appli
857	18	90.0	20	4	US-08-835-281-12	Sequence 12, Appli	930	18	90.0	185	3	US-08-450-582-6	Sequence 6, Appli
858	18	90.0	20	4	US-09-481-593-18	Sequence 18, Appli	931	18	90.0	185	4	US-08-449-731-6	Sequence 6, Appli
859	18	90.0	20	5	PCT-US96-08723-1	Sequence 1, Appli	932	18	90.0	185	4	US-09-198-452A-858	Sequence 858, App
860	18	90.0	22	2	US-08-907-673-1	Sequence 1, Appli	933	18	90.0	187	4	US-09-134-001C-3419	Sequence 3419, Ap
861	18	90.0	22	3	US-08-809-397-6	Sequence 6, Appli	934	18	90.0	189	4	US-09-134-000C-4845	Sequence 4845, Ap
862	18	90.0	22	3	US-09-425-597-6	Sequence 6, Appli	935	18	90.0	199	1	US-08-602-262-2	Sequence 2, Appli
863	18	90.0	22	5	PCT-US95-12502-6	Sequence 6, Appli	936	18	90.0	199	3	US-09-004-716-2	Sequence 2, Appli
864	18	90.0	27	4	US-09-039-780A-82	Sequence 82, Appli	937	18	90.0	199	4	US-09-477-510-2	Sequence 2, Appli
865	18	90.0	28	4	US-09-039-780A-81	Sequence 81, Appli	938	18	90.0	199	4	US-09-543-681A-5219	Sequence 5219, Ap
866	18	90.0	30	4	US-09-039-780A-104	Sequence 104, App	939	18	90.0	199	5	PCT-US94-02889-2	Sequence 2, Appli
867	18	90.0	38	4	US-09-039-780A-83	Sequence 83, Appli	940	18	90.0	201	1	US-07-929-580B-5	Sequence 5, Appli
868	18	90.0	39	4	US-08-939-780A-84	Sequence 84, Appli	941	18	90.0	205	4	US-09-107-532A-5214	Sequence 5214, Ap
869	18	90.0	39	2	US-08-989-291-15	Sequence 15, Appli	942	18	90.0	208	1	US-08-631-607-3	Sequence 3, Appli
870	18	90.0	39	3	US-09-088-244-15	Sequence 15, Appli	943	18	90.0	208	4	US-09-098-358B-3	Sequence 3, Appli
871	18	90.0	39	4	US-09-375-314-15	Sequence 15, Appli	944	18	90.0	210	1	US-07-741-940-4	Sequence 4, Appli
872	18	90.0	39	4	US-09-767-395-15	Sequence 15, Appli	945	18	90.0	210	1	US-08-289-548A-4	Sequence 4, Appli
873	18	90.0	43	4	US-09-039-780A-89	Sequence 89, Appli	946	18	90.0	210	1	US-08-452-654-4	Sequence 4, Appli
874	18	90.0	44	4	US-09-039-780A-88	Sequence 88, Appli	947	18	90.0	210	1	US-08-452-655B-4	Sequence 4, Appli
875	18	90.0	44	4	US-09-039-780A-98	Sequence 98, Appli	948	18	90.0	210	3	US-08-450-582-4	Sequence 4, Appli
876	18	90.0	44	4	US-09-039-780A-99	Sequence 99, Appli	949	18	90.0	210	4	US-08-449-731-4	Sequence 4, Appli
877	18	90.0	56	4	US-09-039-780A-93	Sequence 93, Appli	950	18	90.0	211	4	US-09-328-352-5256	Sequence 5256, Ap
878	18	90.0	60	4	US-09-134-000C-4258	Sequence 4258, Ap	951	18	90.0	214	4	US-09-795-926-16	Sequence 16, Appli
879	18	90.0	61	4	US-09-107-532A-4993	Sequence 4993, Ap	952	18	90.0	220	4	US-09-107-532A-6803	Sequence 6803, Ap
880	18	90.0	66	4	US-09-543-681A-5898	Sequence 5898, Ap	953	18	90.0	223	1	US-07-708-885B-3	Sequence 3, Appli
881	18	90.0	69	4	US-09-328-352-5719	Sequence 5719, Ap	954	18	90.0	223	1	US-07-714-386-3	Sequence 3, Appli
882	18	90.0	88	4	US-09-732-210-1423	Sequence 1423, Ap	955	18	90.0	223	1	US-07-708-888A-3	Sequence 3, Appli
883	18	90.0	90	4	US-09-393-634-76	Sequence 76, Appli	956	18	90.0	229	4	US-09-489-039A-9391	Sequence 9391, Ap
884	18	90.0	101	4	US-09-107-532A-6901	Sequence 6901, Ap	957	18	90.0	232	4	US-09-134-000C-3556	Sequence 3556, Ap
885	18	90.0	105	4	US-09-489-039A-10362	Sequence 10362, Ap	958	18	90.0	235	3	US-09-390-721-4	Sequence 4, Appli
886	18	90.0	113	4	US-09-614-912-186	Sequence 186, App	959	18	90.0	237	3	US-09-248-335-48	Sequence 48, Appli
887	18	90.0	116	4	US-09-489-039A-12512	Sequence 12512, A	960	18	90.0	243	4	US-09-134-000C-3753	Sequence 3753, Ap
888	18	90.0	118	4	US-09-540-236-3531	Sequence 3531, Ap	961	18	90.0	245	4	US-09-543-681A-5699	Sequence 5699, Ap
889	18	90.0	122	4	US-09-540-236-3230	Sequence 3230, Ap	962	18	90.0	247	4	US-09-328-352-7311	Sequence 7311, Ap
890	18	90.0	131	4	US-09-328-352-4806	Sequence 4806, Ap	963	18	90.0	249	4	US-09-252-991A-32233	Sequence 32233, A
891	18	90.0	132	4	US-09-134-000C-3642	Sequence 3642, Ap	964	18	90.0	251	4	US-09-548-004-20	Sequence 20, Appli
892	18	90.0	136	4	US-09-732-210-211	Sequence 211, App	965	18	90.0	251	4	US-09-286-981B-4	Sequence 4, Appli
893	18	90.0	142	1	US-07-951-715A-24	Sequence 24, Appli	966	18	90.0	252	4	US-09-489-039A-13690	Sequence 13690, A
894	18	90.0	142	2	US-08-459-448A-24	Sequence 24, Appli	967	18	90.0	254	4	US-09-198-452A-632	Sequence 632, App
895	18	90.0	142	3	US-08-459-595A-24	Sequence 24, Appli	968	18	90.0	257	2	US-08-637-759B-92	Sequence 92, Appli
896	18	90.0	142	3	US-08-459-504B-24	Sequence 24, Appli	969	18	90.0	257	3	US-08-871-355A-92	Sequence 92, Appli
897	18	90.0	142	3	US-08-459-444-24	Sequence 24, Appli	970	18	90.0	257	4	US-09-201-945-92	Sequence 92, Appli
898	18	90.0	142	3	US-09-547-422-24	Sequence 24, Appli	971	18	90.0	262	1	US-08-622-353-9	Sequence 9, Appli
899	18	90.0	143	3	US-08-946-329A-52	Sequence 52, Appli	972	18	90.0	262	2	US-08-622-352A-11	Sequence 11, Appli
900	18	90.0	143	4	US-09-328-352-7008	Sequence 7008, Ap	973	18	90.0	262	3	US-08-926-390-11	Sequence 11, Appli
901	18	90.0	143	4	US-09-328-352-7008	Sequence 7008, Ap	974	18	90.0	265	4	US-08-795-926-6	Sequence 6, Appli
902	18	90.0	144	4	US-09-543-681A-5427	Sequence 5427, Ap	975	18	90.0	265	4	US-09-134-000C-5370	Sequence 5370, Ap
903	18	90.0	145	3	US-08-720-625-5	Sequence 5, Appli	976	18	90.0	267	4	US-09-543-681A-6389	Sequence 6389, Ap





```

STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/197,484
FILING DATE: 16-FEB-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/935,811
FILING DATE: 26-AUG-1992
APPLICATION DATA:
APPLICATION NUMBER: US 07/874,491
FILING DATE: 27-APR-1992
APPLICATION DATA:
APPLICATION NUMBER: US 07/749,568
FILING DATE: 29-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14137-26-4
TELEPHONE: (206) 467-9600
TELEFAX: (206) 623-6793
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
S-08-197-484-86

Query Match 95.0%; Score 19; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5
b 2 VAEF 5

RESULT 4
S-08-197-484-145
Sequence 145, Application US/08197484
Patent No. 6419931
GENERAL INFORMATION:
APPLICANT: VITIELLO, Maria A.
APPLICANT: CHESTNUT, Robert W.
APPLICANT: SETTE, Alessandro D.
APPLICANT: CELIS, Esteban
APPLICANT: GRAY, Howard
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend Kourie and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/197,484
APPLICATION NUMBER: US/08/197,484
FILING DATE: 16-FEB-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/935,811
FILING DATE: 26-AUG-1992
APPLICATION DATA:
APPLICATION NUMBER: US 07/874,491
FILING DATE: 27-APR-1992
APPLICATION DATA:
APPLICATION NUMBER: US 07/749,568
FILING DATE: 29-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14137-26-4
TELEPHONE: (206) 467-9600
TELEFAX: (206) 623-6793
INFORMATION FOR SEQ ID NO: 145:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-197-484-145

Query Match 95.0%; Score 19; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAEF 5
DB 2 VAEF 5

RESULT 5
US-09-724-566A-73
Sequence 73, Application US/09724566A
Patent No. 6627739
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Basi, Gurigbal
APPLICANT: Doane, Minh Tam
APPLICANT: Frigon, No. 6627739mand
APPLICANT: John, Varghese
APPLICANT: Power, Michael
APPLICANT: Sinha, Sukanto
APPLICANT: Tatsuno, Gwen
APPLICANT: Tung, Jay
APPLICANT: Wang, Shuwen
APPLICANT: McConlogue, Lisa
TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
FILE REFERENCE: 228-US-NEWC2
CURRENT APPLICATION NUMBER: US/09/724,566A
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 09/501,708
PRIOR FILING DATE: 2000-02-10
PRIOR APPLICATION NUMBER: 60/119,571
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/139,172
PRIOR FILING DATE: 1999-06-15
NUMBER OF SEQ ID NOS: 104
SOFTWARE: Fast-SEQ for Windows Version 4.0
SEQ ID NO 73
LENGTH: 9

```



Qy 2 VAEF 5

Query Match	95.0%;	Score 19;	DB 5;	Length 9;
Best Local Similarity	100.0%;	Pred. No. 3e+05;		
Matches 4:	Conservative	0;	Mismatches	0;
	Indels	0;	Gaps	0;

RESULT 8  
US-09-724-566A-72  
; Sequence 72, Application US/09724566A  
; Patent No. 6627739  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, John P.  
; APPLICANT: Basi, Gurigbal  
; APPLICANT: Deane, Minh Tam  
; APPLICANT: Frigon, No. 6627739mand  
; APPLICANT: Jom, Varghese  
; APPLICANT: Power, Michael  
; APPLICANT: Sinha, Sukanto  
; APPLICANT: Tatsuno, Gwen

APPLICANT: Tung, Jay  
APPLICANT: Wang, Shuwen  
APPLICANT: McConlogue, Lisa  
TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and  
TITLE OF INVENTION: Methods  
FILE REFERENCE: 228-US-NEWC2  
CURRENT APPLICATION NUMBER: US/09/724,566A  
CURRENT FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: US 09/501,708  
PRIOR FILING DATE: 2000-02-10  
PRIOR APPLICATION NUMBER: 60/119,571  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: 60/139,172  
PRIOR FILING DATE: 1999-06-15  
NUMBER OF SEQ ID NOS: 104  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 72  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: P10-P4'staD-V peptide inhibitor  
NAME/KEY: MOD\_RES  
LOCATION: 10\_RES  
OTHER INFORMATION: Xaa is statine moiety  
S-09-724-566A-72

Query Match 95.0%; Score 19; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
b 11 VAEF 14

ESULT 9  
S-09-724-566A-97  
Sequence 97, Application US/09724566A  
Patent No. 6627739  
GENERAL INFORMATION:  
APPLICANT: Anderson, John P.  
APPLICANT: Basi, Gurigbal  
APPLICANT: Doane, Minh Tam  
APPLICANT: Frigon, No. 6627739mand  
APPLICANT: John, Varghese  
APPLICANT: Power, Michael  
APPLICANT: Sinha, Sukanto  
APPLICANT: Tatsuno, Gwen  
APPLICANT: Tung, Jay  
APPLICANT: Wang, Shuwen  
APPLICANT: McConlogue, Lisa  
TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and  
TITLE OF INVENTION: Methods  
FILE REFERENCE: 228-US-NEWC2  
CURRENT APPLICATION NUMBER: US/09/724,566A  
CURRENT FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: US 09/501,708  
PRIOR FILING DATE: 2000-02-10  
PRIOR APPLICATION NUMBER: 60/119,571  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: 60/139,172  
PRIOR FILING DATE: 1999-06-15  
NUMBER OF SEQ ID NOS: 104  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 97  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: APP-derived fragment P10-P4' (D-V)  
S-09-724-566A-97

Query Match 95.0%; Score 19; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 11 VAEF 14

RESULT 10  
US-08-733-825-3  
Sequence 3, Application US/08733825  
Patent No. 5837839  
GENERAL INFORMATION:  
APPLICANT: Toth, Matthew J.  
APPLICANT: Huwyler, Leslie R.  
TITLE OF INVENTION: Coding Sequences for Nevalonate  
TITLE OF INVENTION: Pyrophosphate Decarboxylase  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSER: No. 5837839artis Patent and Trademark Department  
STREET: 59 Route 10  
CITY: East Hanover  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07936-1080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/733,825  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/005,652  
FILING DATE: 18-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5837839ak, Henry P.  
REGISTRATION NUMBER: 33200  
REFERENCE/DOCKET NUMBER: 4-20615/P1/CGC 1834  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 277-5110  
TELEFAX: (908) 277-4606  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-733-825-3

Query Match 95.0%; Score 19; DB 2; Length 28;  
Best Local Similarity 100.0%; Pred. No. 1e-02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 20 VAEF 23

RESULT 11  
US-08-487-890A-69  
Sequence 69, Application US/08487890A  
Patent No. 5708149  
GENERAL INFORMATION:  
APPLICANT: Loomore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele

APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Mardin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,890A  
FILING DATE: 07-JUN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-487-890A-69

Query Match 95.0%; Score 19; DB 1; Length 35;  
Best Local Similarity 100.0%; Pred.No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 14 VAEF 17

RESULT 12  
US-08-468-763-1  
Sequence 1, Application US/08468763  
Patent No. 5741671  
GENERAL INFORMATION:  
APPLICANT: Agre, Peter C.  
TITLE OF INVENTION: Isolation, Cloning and Expression of  
TITLE OF INVENTION: Transmembrane Water Channel Proteins  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Allegretti  
STREET: 1001 G Street, N.W.  
CITY: Washington, D.C.  
STATE: D.C.  
COUNTRY: US  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,763  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/393,996  
FILING DATE: 24-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Posorske, Laurence H.  
REGISTRATION NUMBER: 34,698  
REFERENCE/DOCKET NUMBER: 1107.48633  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 508-9100  
TELEFAX: 202 508-9239  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-468-763-1

Query Match 95.0%; Score 19; DB 1; Length 35;  
Best Local Similarity 100.0%; Pred.No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 14 VAEF 17

RESULT 13  
US-08-393-996A-1  
Sequence 1, Application US/08393996A  
Patent No. 5858702  
GENERAL INFORMATION:  
APPLICANT: Agre, Peter C.  
TITLE OF INVENTION: Isolation, Cloning and Expression of  
TITLE OF INVENTION: Transmembrane Water Channel Proteins  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Allegretti  
STREET: 1001 G Street, N.W.  
CITY: Washington, D.C.  
STATE: D.C.  
COUNTRY: US  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,996A  
FILING DATE: 24-FEB-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Posorske, Laurence H.  
REGISTRATION NUMBER: 34,698  
REFERENCE/DOCKET NUMBER: 1107.48633  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 508-9100  
TELEFAX: 202 508-9239  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-393-996A-1

Query Match 95.0%; Score 19; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

2Y 2 VAEF 5  
14 VAEF 17

## RESULT 14

US-08-478-435-69

Sequence 69, Application US/08478435

Patent No. 5922323

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena

APPLICANT: Harkness, Robin

APPLICANT: Schryvers, Anthony

APPLICANT: Chong, Pele

APPLICANT: Gray-Owen, Scott

APPLICANT: Yang, Yan-Ping

APPLICANT: Murdin, Andrew

APPLICANT: Klein, Michel

TITLE OF INVENTION: Transferrin Receptor Genes

NUMBER OF SEQUENCES: 147

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim &amp; McBurney

STREET: Suite 701, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/478,435

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/337,483

FILING DATE: 08-NOV-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/175,116

FILING DATE: 29-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,968

FILING DATE: 08-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-462 MIS:VF

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 69:

SEQUENCE CHARACTERISTICS:

LENGTH: 35 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-478-435-69

Query Match 95.0%; Score 19; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

2Y 2 VAEF 5  
14 VAEF 17

Matches 0; Gaps 0;

RESULT 15  
US-08-337-483-69  
Sequence 69, Application US/08337483  
Patent No. 5922562  
GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena

APPLICANT: Harkness, Robin

APPLICANT: Schryvers, Anthony

APPLICANT: Chong, Pele

APPLICANT: Gray-Owen, Scott

APPLICANT: Yang, Yan-Ping

APPLICANT: Murdin, Andrew

APPLICANT: Klein, Michel

TITLE OF INVENTION: Transferrin Receptor Genes

NUMBER OF SEQUENCES: 147

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim &amp; McBurney

STREET: Suite 701, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/337,483

FILING DATE: 08-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-410 MIS:jb

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 69:

SEQUENCE CHARACTERISTICS:

LENGTH: 35 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-337-483-69

Query Match 95.0%; Score 19; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

QY 2 VAEF 5  
14 VAEF 17

## RESULT 16

US-08-478-373-69

Sequence 69, Application US/08478373

Patent No. 5922841

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena

APPLICANT: Harkness, Robin

APPLICANT: Schryvers, Anthony

APPLICANT: Chong, Pele

APPLICANT: Gray-Owen, Scott

APPLICANT: Yang, Yan-Ping

APPLICANT: Murdin, Andrew

APPLICANT: Klein, Michel

TITLE OF INVENTION: Transferrin Receptor Genes

NUMBER OF SEQUENCES: 147

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim &amp; McBurney

```

; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,373
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-463 MIS:VG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-478-373-69

```

```

Query Match          95.0%; Score 19; DB 2; Length 35;
Best Local Similarity 100.0%; Pred.No.1.3e+02;
Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 VAEF 5
Db      14 VAEF 17

```

```

RESULT 17
US-08-474-671-69
; Sequence 69, Application US/08474671
; Patent No. 6008326
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Murdin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,671
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-465 MIS:VG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-474-671-69

Query Match          95.0%; Score 19; DB 3; Length 35;
Best Local Similarity 100.0%; Pred.No.1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
Db      14 VAEF 17

RESULT 18
US-08-483-577A-69
; Sequence 69, Application US/08483577A
; Patent No. 6015688
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Murdin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 160
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,577A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483

```

```

FILING DATE: 08-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,116
FILING DATE: 29-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,968
FILING DATE: 08-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-511
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-483-577A-69

Query Match 95.0%; Score 19; DB 3; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 2 VAEF 5
b 14 VAEF 17

RESULT 19
US-08-897-438-69
Sequence 69, Application US/08897438
Patent No. 6262016
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena
APPLICANT: Harkness, Robin
APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-Ping
APPLICANT: Mordin, Andrew
APPLICANT: Klein, Michel
TITLE OF INVENTION: Transferrin Receptor Genes
NUMBER OF SEQUENCES: 160
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,438
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/483,577
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,483
FILING DATE: 08-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,116
FILING DATE: 29-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,968

```

```

FILING DATE: 08-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-720
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-897-438-69

Query Match 95.0%; Score 19; DB 3; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5
Db 14 VAEF 17

RESULT 20
US-08-637-654-69
Sequence 69, Application US/08637654
Patent No. 6358727
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M
APPLICANT: Harkness, Robin E
APPLICANT: Schryvers, Anthony B
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-Ping
APPLICANT: Mordin, Andrew D
APPLICANT: Klein, Michel H
TITLE OF INVENTION: HAEMOPHILUS TRANSFERRIN RECEPTOR GENES
NUMBER OF SEQUENCES: 147
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,654
FILING DATE: 05-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA94/00616
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-595
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-637-654-69

```

Query Match 95.0%; Score 19; DB 4; Length 35;  
Best Local Similarity 100.0%; Pred.No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 14 VAEF 17

RESULT 21  
US-08-649-518-69  
; Sequence 69, Application US/08649518  
; Patent No. 6361779  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena  
; APPLICANT: Harkness, Robin  
; APPLICANT: Schryvers, Anthony  
; APPLICANT: Chong, Pele  
; APPLICANT: Gray-Owen, Scott  
; APPLICANT: Yang, Yan-Ping  
; APPLICANT: Mordin, Andrew  
; APPLICANT: Klein, Michel  
; TITLE OF INVENTION: Transferrin Receptor Genes  
; NUMBER OF SEQUENCES: 160  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/649,518  
; FILING DATE: 17-MAY-1996  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/337,483  
; FILING DATE: 08-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/483,577  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/175,116  
; FILING DATE: 29-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/148,968  
; FILING DATE: 08-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-608  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-649-518-69

Query Match 95.0%; Score 19; DB 4; Length 35;  
Best Local Similarity 100.0%; Pred.No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

Db 14 VAEF 17

RESULT 22  
US-09-393-634-80  
; Sequence 80, Application US/09393634  
; Patent No. 6558910  
; GENERAL INFORMATION:  
; APPLICANT: Zuker, Charles S.  
; APPLICANT: Adler, Jon Elliot  
; APPLICANT: Ryba, Nick  
; APPLICANT: Mueller, Ken  
; APPLICANT: Hoon, Mark  
; APPLICANT: The Regents of the University of California  
; APPLICANT: as represented by the Secretary of the  
; APPLICANT: Department of Health and Human Services  
; TITLE OF INVENTION: SF, a No. 6558910el Family of Taste Receptors  
; FILE REFERENCE: 02307E-098000US  
; CURRENT APPLICATION NUMBER: US/09/393,634  
; CURRENT FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 92  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 80  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human GR24  
US-09-393-634-80

Query Match 95.0%; Score 19; DB 4; Length 68;  
Best Local Similarity 100.0%; Pred.No. 2.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 14 VAEF 17

RESULT 23  
US-09-489-039A-13555  
; Sequence 13555, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary B'reton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 13555  
; LENGTH: 69  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-13555

Query Match 95.0%; Score 19; DB 4; Length 69;  
Best Local Similarity 100.0%; Pred.No. 2.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 5 VAEF 8

RESULT 24  
US-09-328-352-7111  
; Sequence 7111, Application US/09328352

Patent No. 6562958  
 GENERAL INFORMATION:  
 APPLICANT: Gary L. Breton et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: GTC99-03PA  
 CURRENT APPLICATION NUMBER: US/09/328,352  
 CURRENT FILING DATE: 1999-06-04  
 NUMBER OF SEQ ID NOS: 8252  
 SEQ ID NO 7111  
 LENGTH: 71  
 TYPE: PRT  
 ORGANISM: Acinetobacter baumannii  
 S-09-328-352-7111

Query Match 95.0%; Score 19; DB 4; Length 71;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
 b 26 VAEF 29

RESULT 25  
 S-08-321-071A-10  
 Sequence 10, Application US/08321071A  
 Patent No. 5672686  
 GENERAL INFORMATION:  
 APPLICANT: CHITTENDEN, Thomas D.  
 TITLE OF INVENTION: APOPTOSIS RELATED PROTEIN Bcl-Y, AND METHODS  
 TITLE OF INVENTION: OP USE THEREOF  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hale and Dorr  
 STREET: 1455 Pennsylvania Avenue, N.W.  
 CITY: Washington  
 STATE: D.C.  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/321,071A  
 FILING DATE: 11-OCT-1994  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/10103  
 FILING DATE: 09-AUG-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/287,427  
 FILING DATE: 09-AUG-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: WIXON, HENRY N.  
 REGISTRATION NUMBER: 32,073  
 REFERENCE/DOCKET NUMBER: 104322.121CIP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-942-8400  
 TELEFAX: 202-942-8484  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 73 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 JS-08-321-071A-10

Query Match 95.0%; Score 19; DB 1; Length 73;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 DB 47 VAEF 50

RESULT 26  
 US-08-894-139-10  
 Sequence 10, Application US/08894139  
 Patent No. 6448376  
 GENERAL INFORMATION:  
 APPLICANT: LA THANGUE, NICHOLAS B.  
 APPLICANT: BERNARDS, RENE  
 APPLICANT: HIJMAN, ELEANORE M.  
 TITLE OF INVENTION: TRANSCRIPTION FACTOR B2F-5  
 NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: NIXON & VANDERHYZE P.C.  
 STREET: 1100 NORTH GLEBE ROAD  
 CITY: ARLINGTON  
 STATE: VIRGINIA  
 COUNTRY: U.S.A.  
 ZIP: 22201-4714  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/894,139  
 FILING DATE: 13-AUG-1997  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: WILSON, MARY J.  
 REGISTRATION NUMBER: 32,955  
 REFERENCE/DOCKET NUMBER: 620-22  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 816-4000  
 TELEFAX: (703) 816-4100  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 74 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-894-139-10

Query Match 95.0%; Score 19; DB 4; Length 74;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 DB 23 VAEF 26

RESULT 27  
 US-09-732-210-1412  
 Sequence 1412, Application US/09732210  
 Patent No. 6573361  
 GENERAL INFORMATION:  
 APPLICANT: Bunkers, Greg J.  
 APPLICANT: Liang, Jihong  
 APPLICANT: Mittanck, Cindy A.  
 APPLICANT: Seale, Jeffrey W.  
 APPLICANT: Wu, Yonnie S.  
 TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
 FILE REFERENCE: 38-21(15036)B  
 CURRENT APPLICATION NUMBER: US/09/732,210  
 CURRENT FILING DATE: 2000-12-07  
 PRIOR APPLICATION NUMBER: US 60/169,513  
 PRIOR FILING DATE: 1999-12-07



```
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 1412
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-732-210-1412

Query Match          95.0%; Score 19; DB 4; Length 88;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
DB      11 VAEF 14

RESULT 28
US-07-612-674-12
; Sequence 12, Application US/07612674
; Patent No. 5658792
; GENERAL INFORMATION:
; APPLICANT: NUELLE, MARK J.
; APPLICANT: MCCLUNG, J. KEITH
; APPLICANT: STEWART, DAVID A.
; APPLICANT: DANNER, DAVID B.
; TITLE OF INVENTION: AN ANTIPROLIFERATIVE PROTEIN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARYL & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/612,674
; FILING DATE: 19901114
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: 5683/92332
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3027
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-612-674-12

Query Match          95.0%; Score 19; DB 1; Length 93;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
DB      18 VAEF 21

RESULT 29
US-09-543-681A-5816
; Sequence 5816, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5816
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5816

Query Match          95.0%; Score 19; DB 4; Length 98;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
DB      21 VAEF 24

RESULT 30
US-08-464-517-8
; Sequence 8, Application US/08464517
; Patent No. 5869840
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
```

S-08-464-517-8

Query Match 95.0%; Score 19; DB 2; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 2 VAEF 5  
 ||||  
 b 13 VAEF 16

RESULT 31

S-08-246-361A-8  
 Sequence 8, Application US/08246361A  
 Patent No. 5998582

## GENERAL INFORMATION:

APPLICANT: BEACH, David H.  
 TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
 NUMBER OF SEQUENCES: 50  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 60 State Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII(text)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/246,361A  
 FILING DATE: 19-MAY-1994  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/963,308  
 FILING DATE: 16-OCT-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/888,178  
 FILING DATE: 26-MAY-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/701,514  
 FILING DATE: 16-MAY-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Matthew P. Vincent  
 REGISTRATION NUMBER: 36,709  
 REFERENCE/DOCKET NUMBER: MII-004C  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-7400  
 TELEFAX: (617) 227-5941

## INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 100 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide

US-08-246-361A-8

Query Match 95.0%; Score 19; DB 2; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 ||||  
 DB 13 VAEF 16

RESULT 32

US-08-463-772-8  
 Sequence 8, Application US/08463772  
 Patent No. 6066501

## GENERAL INFORMATION:

APPLICANT: BEACH, David H.  
 TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
 NUMBER OF SEQUENCES: 50  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 60 State Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII(text)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/463,772  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/963,308  
 FILING DATE: 16-OCT-1992  
 APPLICATION NUMBER: US 07/888,178  
 FILING DATE: 26-MAY-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/701,514  
 FILING DATE: 16-MAY-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Matthew P. Vincent  
 REGISTRATION NUMBER: 36,709  
 REFERENCE/DOCKET NUMBER: MII-004C  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-7400  
 TELEFAX: (617) 227-5941

## INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 100 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide

US-08-463-772-8

Query Match 95.0%; Score 19; DB 3; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 ||||  
 DB 13 VAEF 16

RESULT 33

PCT-US93-05000-8

Sequence 8, Application PC/TUS9305000

## GENERAL INFORMATION:

APPLICANT: MITOTIX  
 TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto  
 NUMBER OF SEQUENCES: 42  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Militia Drive  
 CITY: Lexington  
 STATE: Massachusetts  
 COUNTRY: US  
 ZIP: 02173

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/05000  
FILING DATE: 19930525  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/888,178  
FILING DATE: 26-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: CSHL91-02A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 616-861-9540  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCT-US93-05000-8

Query Match 95.0%; Score 19; DB 5; Length 100;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 13 VAEF 16

## RESULT 34

US-08-580-988A-21  
Sequence 21, Application US/08580988A  
Patent No. 5956161  
GENERAL INFORMATION:  
APPLICANT: Aggarwal et al.  
TITLE OF INVENTION: Tumor Necrosis Factor  
TITLE OF INVENTION: Receptor-I-Associated Protein Kinase And Methods  
TITLE OF INVENTION: For Its Use  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dr. Benjamin A. Adler  
STREET: 8011 Candle Lane  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77071

COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 Mb floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh  
SOFTWARE: Microsoft Word for Macintosh  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/580,988A  
FILING DATE: January 3, 1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benjamin Aaron Adler, Ph.D., J.D.  
REGISTRATION NUMBER: 35,423  
REFERENCE/DOCKET NUMBER: D5721C1P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-777-2321  
TELEFAX: 713-777-6908  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 101 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE:

DESCRIPTION: protein  
HYPOTHETICAL: no  
ANTI-SENSE: no  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-08-580-988A-21

Query Match 95.0%; Score 19; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 70 VAEF 73

## RESULT 35

US-09-540-236-3612  
Sequence 3612, Application US/09540236  
Patent No. 6673910  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATA  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540,236  
CURRENT FILING DATE: 2000-04-04  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 3612  
LENGTH: 103  
TYPE: PRT  
ORGANISM: M.catarrhalis  
US-09-540-236-3612

Query Match 95.0%; Score 19; DB 4; Length 103;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 65 VAEF 68

## RESULT 36

US-08-464-517-25  
Sequence 25, Application US/08464517  
Patent No. 5869640  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,517  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
3-08-464-517-25

Query Match 95.0%; Score 19; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 2 VAEF 5  
|||  
66 VAEF 69

3

RESULT 37  
3-08-246-361A-25  
Sequence 25, Application US/08246361A  
Patent No. 5998582  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,361A  
FILING DATE: 19-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
US-08-246-361A-25

Query Match 95.0%; Score 19; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
Db 66 VAEF 69

RESULT 38  
US-08-463-772-25  
Sequence 25, Application US/08463772  
Patent No. 6066501  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,772  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-463-772-25

Query Match 95.0%; Score 19; DB 3; Length 106;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|||  
Db 66 VAEF 69

RESULT 39  
US-09-489-039A-8721  
Sequence 8721, Application US/09489039A  
Patent No. 6610836

```

; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8721
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
JS-09-489-039A-8721

Query Match 95.0%; Score 19; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 43 VAEF 46

RESULT 40
PCT-US93-05000-25
; Sequence 25, Application PC/TUS9305000
; GENERAL INFORMATION:
; APPLICANT: MITOTIX
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
PCT-US93-05000-25

Query Match 95.0%; Score 19; DB 5; Length 106;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 66 VAEF 69

RESULT 41
US-09-489-039A-9915
; Sequence 9915, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9915
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9915

Query Match 95.0%; Score 19; DB 4; Length 115;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 50 VAEF 53

RESULT 42
US-09-252-991A-21614
; Sequence 21614, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21614
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21614

Query Match 95.0%; Score 19; DB 4; Length 122;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5
DB 85 VAEF 88

RESULT 43
US-09-489-039A-7200
; Sequence 7200, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
```

PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 7200  
LENGTH: 123  
TYPE: PRT

ORGANISM: Klebsiella pneumoniae  
S-09-489-039A-7200

Query Match 95.0%; Score 19; DB 4; Length 123;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
b 61 VAEF 64

RESULT 44

S-08-529-055-53  
Sequence 53, Application US/08529055  
Patent No. 6592876

GENERAL INFORMATION:

APPLICANT: Briles, David E.

APPLICANT: McDaniel, Larry S.

APPLICANT: Swiatlo, Edwin

APPLICANT: Yoether, Janet

APPLICANT: Brooks-Walter, Alexis

TITLE OF INVENTION: Pneumococcal Genes, Portions

TITLE OF INVENTION: Thereof, Expression Products

TITLE OF INVENTION: Thereof, and Uses of Such Genes,

TITLE OF INVENTION: Portions and Products

NUMBER OF SEQUENCES: 73

CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtis, Morris & Safford, P.C.

STREET: 530 Fifth Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/529,055

FILING DATE: 15-SEP-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Frommer, William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454312-2400

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 53:

SEQUENCE CHARACTERISTICS:

LENGTH: 129 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

S-08-529-055-53

Query Match 95.0%; Score 19; DB 4; Length 129;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
b 7 VAEF 10

RESULT 45

US-09-489-039A-12835  
Sequence 12835, Application US/09489039A

Patent No. 6610836

GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.2004001

CURRENT APPLICATION NUMBER: US/09/489,039A

CURRENT FILING DATE: 2000-01-27

PRIOR FILING DATE: US 60/117,747

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 12835

LENGTH: 143

TYPE: PRT

ORGANISM: Klebsiella pneumoniae

US-09-489-039A-12835

Query Match 95.0%; Score 19; DB 4; Length 143;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 138 VAEF 141

RESULT 46

US-09-198-452A-167

Sequence 167, Application US/09198452A

Patent No. 6559294

GENERAL INFORMATION:

APPLICANT: Griffois, R.

TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments

TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention

TITLE OF INVENTION: and treatment of infection

FILE REFERENCE: 9710-003-999

CURRENT APPLICATION NUMBER: US/09/198,452A

CURRENT FILING DATE: 1998-11-24

NUMBER OF SEQ ID NOS: 6849

SEQ ID NO 167

LENGTH: 145

TYPE: PRT

ORGANISM: Chlamydia pneumoniae

US-09-198-452A-167

Query Match 95.0%; Score 19; DB 4; Length 145;

Best Local Similarity 100.0%; Pred. No. 5.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 31 VAEF 34

RESULT 47

US-09-134-000C-4292

Sequence 4292, Application US/09134000C

Patent No. 6617156

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 032796-032

CURRENT APPLICATION NUMBER: US/09/134,000C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/055,778

PRIOR FILING DATE: 1997-08-15

NUMBER OF SEQ ID NOS: 6812

SOFTWARE: PatentIn version 3.1

SEQ ID NO 4292  
LENGTH: 145  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-4292

Query Match 95.0%; Score 19; DB 4; Length 145;  
Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 113 VAEF 116

RESULT 48  
US-08-460-694-3

; Sequence 3, Application US/08460694  
; Patent No. 5858655  
; GENERAL INFORMATION:  
; APPLICANT: Arnold, Andrew  
; TITLE OF INVENTION: PRAD1 Cyclin and its cDNA  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,694  
FILING DATE: 02-JUN-1995

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070002

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 150 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-460-694-3

Query Match 95.0%; Score 19; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 5.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 66 VAEF 69

RESULT 49

US-08-460-744-3

; Sequence 3, Application US/08460744

; Patent No. 6107541

; GENERAL INFORMATION:

; APPLICANT: Arnold, Andrew

; TITLE OF INVENTION: PRAD1 Cyclin and its cDNA

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/460,744

FILING DATE: 02-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McConathy, Evelyn H.

REGISTRATION NUMBER: 35,279

REFERENCE/DOCKET NUMBER: 0609.4070005

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 150 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-460-744-3

Query Match 95.0%; Score 19; DB 3; Length 150;  
Best Local Similarity 100.0%; Pred. No. 5.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 66 VAEF 69

RESULT 50

US-07-667-711B-3

; Sequence 3, Application US/07667711B

; Patent No. 6110700

; GENERAL INFORMATION:

; APPLICANT: ARNOLD, ANDREW

; TITLE OF INVENTION: Prad1 Cyclin and its cDNA

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

; STREET: 1100 NEW YORK AVE., NW, SUITE 600

; CITY: WASHINGTON

; STATE: DC

; COUNTRY: USA

; ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/667,711B

FILING DATE: 11-MAR-1991

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MCPHAIL, DONALD R.

REGISTRATION NUMBER: 35,811

REFERENCE/DOCKET NUMBER: 0609.4070000

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

```

LENGTH: 150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
3-07-667-711B-3

Query Match          95.0%; Score 19; DB 3; Length 150;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y      2 VAEF 5
      ||||
      66 VAEF 69

RESULT 51
3-08-679-493A-208
Sequence 208, Application US/08679493A
Patent No. 6303295
GENERAL INFORMATION:
APPLICANT: Taylor, Ethan W.
TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
FILE REFERENCE: 55-95
CURRENT APPLICATION NUMBER: US/08/679,493A
PRIOR FILING DATE: 1996-07-12
PRIOR APPLICATION NUMBER: 60/001203
PRIOR FILING DATE: 1995-07-14
PRIOR APPLICATION NUMBER: 60/003,112
PRIOR FILING DATE: 1995-09-01
NUMBER OF SEQ ID NOS: 216
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 208
LENGTH: 152
TYPE: PRT
ORGANISM: blueshark
3-08-679-493A-208

Query Match          95.0%; Score 19; DB 4; Length 152;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y      2 VAEF 5
      ||||
      92 VAEF 95

RESULT 52
3-09-134-000C-4087
Sequence 4087, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4087
LENGTH: 153
TYPE: PRT
ORGANISM: Enterococcus faecalis
3-09-134-000C-4087

Query Match          95.0%; Score 19; DB 4; Length 153;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y      2 VAEF 5

us-09-594-978a-1.ra1

Db      101 VAEF 104
      ||||

RESULT 53
US-09-252-991A-20805
Sequence 20805, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 20805
LENGTH: 156
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20805

Query Match          95.0%; Score 19; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY     2 VAEF 5
      ||||
      84 VAEF 87

Db      100 VAEF 103
      ||||

RESULT 54
US-09-370-838-203
Sequence 203, Application US/09370838
Patent No. 644425
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Mohamath, Roadoh
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
FILE REFERENCE: 210121.475C1
CURRENT APPLICATION NUMBER: US/09/370,838
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/285,323
EARLIER FILING DATE: 1999-04-02
NUMBER OF SEQ ID NOS: 289
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 203
LENGTH: 164
TYPE: PRT
ORGANISM: Homo sapien
US-09-370-838-203

Query Match          95.0%; Score 19; DB 4; Length 164;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY     2 VAEF 5
      ||||
      100 VAEF 103

Db      100 VAEF 103
      ||||

RESULT 55
US-09-540-236-2222
Sequence 2222, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:
```



APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540,236  
CURRENT FILING DATE: 2000-04-04  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 2222  
LENGTH: 166  
TYPE: PRT  
ORGANISM: M.cattarrhalis  
US-09-540-236-2222

Query Match 95.0%; Score 19; DB 4; Length 166;  
Best Local Similarity 100.0%; Pred. No. 6.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 97 VAEF 100

RESULT 56  
US-08-471-058-16  
Sequence 16, Application US/08471058  
Patent No. 5770443  
GENERAL INFORMATION:  
APPLICANT: Kiefer, Michael C.  
TITLE OF INVENTION: NOVEL APOPTOSIS MODULATING  
TITLE OF INVENTION: PROTEINS, DNA ENCODING THE PROTEINS AND METHODS OF USE  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,058  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/320,157  
FILING DATE: 07-OCT-1994  
APPLICATION NUMBER: 08/160,067  
FILING DATE: 30-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lehnhardt, Susan K  
REGISTRATION NUMBER: 33,943  
REFERENCE/DOCKET NUMBER: 23647-20007.12  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-813-5600  
TELEFAX: 415-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 172 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-471-058-16

Query Match 95.0%; Score 19; DB 1; Length 172;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 122 VAEF 125

RESULT 57  
US-08-690-095-3  
Sequence 3, Application US/08690095  
Patent No. 5792648  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Au-Yang, Janice  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN MACROPHAGE ANTIGEN  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/690,095  
FILING DATE: Filed Herewith  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0110 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 172 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 293274  
US-08-690-095-3

Query Match 95.0%; Score 19; DB 1; Length 172;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 122 VAEF 125

RESULT 58  
US-08-471-057-16  
Sequence 16, Application US/08471057  
Patent No. 6015687  
GENERAL INFORMATION:  
APPLICANT: KIEFER, MICHAEL C.  
APPLICANT: BARR, PHILIP J.  
TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA  
ENCODING THE PROTEINS AND METHODS OF USE THEREOF  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto

STATE: California  
 COUNTRY: USA  
 ZIP: 94304-1018  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/471,057  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/320,157  
 FILING DATE: 07-OCT-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: LEHNHARDT, SUSAN K.  
 REGISTRATION NUMBER: 33,943  
 REFERENCE/DOCKET NUMBER: 23647-20007.20  
 TELEPHONE: (415) 813-5600  
 TELEFAX: (415) 494-0792  
 TELEX: 706141  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 172 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 S-08-471-057-16

Query Match 95.0%; Score 19; DB 3; Length 172;  
 Best Local Similarity 100.0%; Pred.No. 6.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
 b 122 VAEF 125

RESULT 59  
 S-09-113-789-3  
 Sequence 3, Application US/09113789  
 Patent No. 6034219  
 GENERAL INFORMATION:  
 APPLICANT: Hillman, Jennifer L.  
 APPLICANT: Au-Young, Jarice  
 APPLICANT: Goli, Surya K.  
 TITLE OF INVENTION: NOVEL HUMAN MACROPHAGE ANTIGEN  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: U.S.  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/113,789  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/690,095  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ballings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PF-0110 US  
 TELECOMMUNICATION INFORMATION:

STATE: California  
 COUNTRY: USA  
 ZIP: 94304-1018  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/471,057  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/320,157  
 FILING DATE: 07-OCT-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: LEHNHARDT, SUSAN K.  
 REGISTRATION NUMBER: 33,943  
 REFERENCE/DOCKET NUMBER: 23647-20007.20  
 TELEPHONE: (415) 813-5600  
 TELEFAX: (415) 494-0792  
 TELEX: 706141  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 172 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 S-08-471-057-16

Query Match 95.0%; Score 19; DB 3; Length 172;  
 Best Local Similarity 100.0%; Pred.No. 6.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
 Db 122 VAEF 125

RESULT 60  
 US-08-470-865-16  
 ; Sequence 16, Application US/08470865  
 ; Patent No. 6586395  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KIEFER, MICHAEL C.  
 ; APPLICANT: BARR, PHILIP J.  
 ; TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA  
 ; TITLE OF INVENTION: ENCODING THE PROTEINS AND METHODS OF USE THEREOF  
 ; NUMBER OF SEQUENCES: 22  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MORRISON & FOERSTER  
 ; STREET: 755 Page Mill Road  
 ; CITY: Palo Alto  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94304-1018  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/470,865  
 ; FILING DATE: 06-JUN-1995  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/320,157  
 ; FILING DATE: 07-OCT-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: LEHNHARDT, SUSAN K.  
 ; REGISTRATION NUMBER: 33,943  
 ; REFERENCE/DOCKET NUMBER: 23647-20007.20  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 813-5600  
 ; TELEFAX: (415) 494-0792  
 ; TELEX: 706141  
 ; INFORMATION FOR SEQ ID NO: 16:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 172 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-470-865-16

Query Match 95.0%; Score 19; DB 4; Length 172;  
 Best Local Similarity 100.0%; Pred.No. 6.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5

DB 122 VAEF 125

RESULT 61  
US-08-193-977-4  
; Sequence 4, Application US/08193977  
; Patent No. 5625031  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, KEVIN R.  
; APPLICANT: COLEMAN, KEVIN G.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND  
; TITLE OF INVENTION: P34CDC2 CELL CYCLE REGULATORY KINASES AND HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS E7 ONCOPROTEIN  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: REED & ROBINS  
; STREET: 635 BRYANT STREET  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/193,977  
; FILING DATE: 08-FEB-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROBINS, ROBERTA L.  
; REGISTRATION NUMBER: 33,208  
; REFERENCE/DOCKET NUMBER: 5998-0016  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 617-8999  
; TELEFAX: (415) 327-3231  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 173 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-193-977-4

Query Match 95.0%; Score 19; DB 1; Length 173;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 100 VAEF 103

RESULT 62  
US-09-489-039A-9869  
; Sequence 3669, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709-2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 9869  
; LENGTH: 174  
; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9869

Query Match 95.0%; Score 19; DB 4; Length 174;  
Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 77 VAEF 80

RESULT 63  
US-08-737-980-2  
; Sequence 2, Application US/08737980  
; Patent No. 5843773  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Apoptosis Regulating Gene  
; NUMBER OF SEQUENCES: 2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/737,980  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: KR 1995-6266  
; FILING DATE: 24-MAR-1995  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 175 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-737-980-2

Query Match 95.0%; Score 19; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 122 VAEF 125

RESULT 64  
US-09-328-352-6739  
; Sequence 6739, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6739  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6739

Query Match 95.0%; Score 19; DB 4; Length 180;  
Best Local Similarity 100.0%; Pred. No. 7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 8 VAEF 11

RESULT 65  
S-09-252-991A-24341  
Sequence 24341, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 24341  
LENGTH: 185  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
S-09-252-991A-24341  
Query Match 95.0%; Score 19; DB 4; Length 185;  
Best Local Similarity 100.0%; Pred. No. 7.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Y 2 VAEF 5  
b 4 VAEF 7  
RESULT 66  
S-09-543-681A-7313  
Sequence 7313, Application US/09543681A  
Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543,681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 7313  
LENGTH: 189  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
S-09-543-681A-7313  
Query Match 95.0%; Score 19; DB 4; Length 189;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Y 2 VAEF 5  
b 73 VAEF 76  
RESULT 67  
S-08-816-241-1  
Sequence 1, Application US/08816241  
Patent No. 5804185  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
TITLE OF INVENTION: NOVEL RNA EDITING ENZYME  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
S-09-594-978a-1.rai

STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/816,241  
FILING DATE: Filed Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0239 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-885-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 190 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PROSUT09  
CLONE: 1646823  
US-08-816-241-1  
Query Match 95.0%; Score 19; DB 1; Length 190;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
Db 104 VAEF 107  
RESULT 68  
US-09-128-395-1  
Sequence 1, Application US/09128395  
Patent No. 6087108  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL RNA EDITING ENZYME  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/128,395  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/816,241  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0239 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 190 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PROSTUT09  
; CLONE: 1646823  
US-09-128-395-1

Query Match 95.0%; Score 19; DB 3; Length 190;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 104 VAEF 107

RESULT 69  
US-09-134-000C-3800  
; Sequence 3800, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3800  
; LENGTH: 194  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-3800

Query Match 95.0%; Score 19; DB 4; Length 194;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 115 VAEF 118

RESULT 70  
US-09-232-200-85  
; Sequence 85, Application US/09232200A  
; Patent No. 6288213  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WHI97-21p3MB  
; CURRENT APPLICATION NUMBER: US/09/232,200A  
; CURRENT FILING DATE: 1999-01-14  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491

; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 85  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Aspergillus nidulans  
US-09-232-200-85

Query Match 95.0%; Score 19; DB 3; Length 199;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 99 VAEF 102

RESULT 71  
US-09-232-197-85  
; Sequence 85, Application US/09232197A  
; Patent No. 6300096  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WHI97-21p3MA  
; CURRENT APPLICATION NUMBER: US/09/232,197A  
; CURRENT FILING DATE: 1999-01-14  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 85  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Aspergillus nidulans  
US-09-232-197-85

Query Match 95.0%; Score 19; DB 4; Length 199;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 99 VAEF 102

RESULT 72  
US-09-232-201-85  
; Sequence 85, Application US/09232201A  
; Patent No. 6348321  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WHI97-21p3MC  
; CURRENT APPLICATION NUMBER: US/09/232,201A  
; CURRENT FILING DATE: 1999-01-14  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15

EARLIER APPLICATION NUMBER: 60/093,491  
EARLIER FILING DATE: 1998-07-20  
EARLIER APPLICATION NUMBER: 60/110,941  
EARLIER FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 85  
LENGTH: 199  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
3-09-232-201-85

Query Match 95.0%; Score 19; DB 4; Length 199;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
99 VAEF 102

RESULT 73  
3-09-232-195-85  
Sequence 85, Application US/09232195A  
Patent No. 6657049  
GENERAL INFORMATION:  
APPLICANT: Stahl, Andreas  
APPLICANT: Hirsch, David J.  
APPLICANT: Lodish, Harvey F.  
APPLICANT: Gimeno, Ruth E.  
APPLICANT: Tartaglia, Louis A.  
TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
FILE REFERENCE: WHI97-21P3MD  
CURRENT APPLICATION NUMBER: US/09/232,195A  
CURRENT FILING DATE: 1999-01-04  
EARLIER APPLICATION NUMBER: 60/071,374  
EARLIER FILING DATE: 1998-01-15  
EARLIER APPLICATION NUMBER: 60/093,491  
EARLIER FILING DATE: 1998-07-20  
EARLIER APPLICATION NUMBER: 60/110,941  
EARLIER FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 85  
LENGTH: 199  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
S-09-232-195-85

Query Match 95.0%; Score 19; DB 4; Length 199;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
99 VAEF 102

RESULT 74  
S-09-107-532A-4612  
Sequence 4612, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA

ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 4612:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 203 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (5) LOCATION 1...203  
SEQUENCE DESCRIPTION: SEQ ID NO: 4612:  
US-09-107-532A-4612

Query Match 95.0%; Score 19; DB 4; Length 203;  
Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
63 VAEF 66

RESULT 75  
US-08-684-024-1  
Sequence 1, Application US/08684024  
Patent No. 5834298  
GENERAL INFORMATION:  
APPLICANT: Benezra, Robert  
TITLE OF INVENTION: GENE ENCODING THE HUMAN HOMOLOG OF MAD2  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/684,024  
FILING DATE: 19-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 1747/46621-A

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0526

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 205 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

IS-08-684-024-1

Query Match 95.0%; Score 19; DB 2; Length 205;

Best Local Similarity 100.0%; Pred. No. 8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

&gt; 2 VAEF 5

20 VAEF 23

Search completed: May 24, 2004, 17:43:20  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

M protein - protein search, using sw model

run on: May 24, 2004, 17:34:23 ; Search time 39.6429 Seconds  
(without alignments)  
35.179 Million cell updates/sec

title: \* US-09-594-978A-1

erfect score: 20

sequence: 1 XVAEP 5

coring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 1149313 segs, 278921704 residues

total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Published Applications AA:

1: /cgn2\_6/prodata/1/pubaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/1/pubaa/US07\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/1/pubaa/US07\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/1/pubaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/1/pubaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/1/pubaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/1/pubaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/1/pubaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/1/pubaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/1/pubaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/1/pubaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/prodata/1/pubaa/US09\_PUB.pep.\*  
13: /cgn2\_6/prodata/1/pubaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/1/pubaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/1/pubaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/1/pubaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/prodata/1/pubaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/prodata/1/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	19	95.0	8	9	US-09-791-378-343
2	19	95.0	8	10	US-09-791-393-206
3	19	95.0	8	10	US-09-791-393-206
4	19	95.0	9	14	US-10-128-711-86
5	19	95.0	9	14	US-10-128-711-145
6	19	95.0	10	12	US-09-908-943A-197
7	19	95.0	12	12	US-09-908-943A-196
8	19	95.0	12	14	US-10-032-818-34
9	19	95.0	13	14	US-10-239-313A-379
10	19	95.0	15	14	US-10-119-528-7
11	19	95.0	20	10	US-09-171-432A-62
12	19	95.0	20	10	US-09-171-432A-63
13	19	95.0	25	10	US-09-171-432A-65
14	19	95.0	28	9	US-09-864-761-47201
15	19	95.0	28	12	US-10-424-599-216297

16	19	95.0	35	14	US-10-043-344-69	Sequence 59, Appl
17	19	95.0	37	15	US-10-242-355-570	Sequence 570, Appl
18	19	95.0	37	15	US-10-430-752A-49	Sequence 49, Appl
19	19	95.0	38	15	US-10-430-752A-45	Sequence 45, Appl
20	19	95.0	41	15	US-10-372-003A-71	Sequence 71, Appl
21	19	95.0	43	9	US-09-864-761-37202	Sequence 37202, A
22	19	95.0	43	12	US-10-424-599-183032	Sequence 183032, A
23	19	95.0	45	12	US-10-424-599-171318	Sequence 171318, A
24	19	95.0	48	9	US-09-864-761-38166	Sequence 38166, A
25	19	95.0	52	12	US-10-424-599-205632	Sequence 205632, A
26	19	95.0	55	12	US-10-424-599-192247	Sequence 192247, A
27	19	95.0	55	12	US-10-424-599-233502	Sequence 233502, A
28	19	95.0	55	15	US-10-372-003A-63	Sequence 63, Appl
29	19	95.0	55	15	US-10-372-003A-67	Sequence 67, Appl
30	19	95.0	56	14	US-10-219-329-4	Sequence 4, Appl
31	19	95.0	56	14	US-10-153-185-4	Sequence 4, Appl
32	19	95.0	56	14	US-10-219-561-4	Sequence 4, Appl
33	19	95.0	57	12	US-10-424-599-215734	Sequence 215734, A
34	19	95.0	57	12	US-10-424-599-260762	Sequence 260762, A
35	19	95.0	57	15	US-10-372-003A-75	Sequence 75, Appl
36	19	95.0	58	12	US-10-424-599-214930	Sequence 214930, A
37	19	95.0	60	9	US-09-864-761-45291	Sequence 45291, A
38	19	95.0	68	9	US-09-393-634-80	Sequence 80, Appl
39	19	95.0	68	11	US-09-864-408A-40	Sequence 40, Appl
40	19	95.0	68	12	US-10-364-861-80	Sequence 80, Appl
41	19	95.0	68	14	US-10-383-982-80	Sequence 80, Appl
42	19	95.0	69	9	US-09-864-761-45716	Sequence 45716, A
43	19	95.0	69	9	US-09-738-626-5596	Sequence 5596, Ap
44	19	95.0	69	12	US-10-424-599-185385	Sequence 185385, A
45	19	95.0	70	14	US-10-029-386-28645	Sequence 28645, A
46	19	95.0	71	12	US-10-424-599-202605	Sequence 202605, A
47	19	95.0	71	12	US-10-424-599-274981	Sequence 274981, A
48	19	95.0	71	14	US-10-029-386-30348	Sequence 30348, A
49	19	95.0	72	14	US-10-029-386-29320	Sequence 29320, A
50	19	95.0	74	14	US-10-214-188-10	Sequence 10, Appl
51	19	95.0	74	14	US-10-156-761-9071	Sequence 9071, Ap
52	19	95.0	75	12	US-10-424-599-156128	Sequence 156128, A
53	19	95.0	75	12	US-10-424-599-261424	Sequence 261424, A
54	19	95.0	76	12	US-10-424-599-253204	Sequence 253204, A
55	19	95.0	77	12	US-10-424-599-211908	Sequence 211908, A
56	19	95.0	77	14	US-10-335-977-7493	Sequence 7493, Ap
57	19	95.0	77	14	US-10-029-386-32709	Sequence 32709, A
58	19	95.0	78	12	US-10-424-599-162039	Sequence 162039, A
59	19	95.0	78	12	US-10-424-599-262701	Sequence 262701, A
60	19	95.0	79	12	US-10-363-616-368	Sequence 368, App
61	19	95.0	80	12	US-10-424-599-216097	Sequence 216097, A
62	19	95.0	82	15	US-10-367-980A-14	Sequence 14, Appl
63	19	95.0	84	11	US-09-864-408A-5282	Sequence 5282, Ap
64	19	95.0	84	14	US-10-078-090-125	Sequence 125, App
65	19	95.0	85	12	US-10-424-599-155636	Sequence 155636, A
66	19	95.0	85	12	US-10-424-599-201853	Sequence 201853, A
67	19	95.0	89	9	US-09-815-242-11224	Sequence 11224, A
68	19	95.0	89	9	US-09-815-242-11245	Sequence 11245, A
69	19	95.0	89	12	US-10-282-122A-58489	Sequence 58489, A
70	19	95.0	89	12	US-10-282-122A-58530	Sequence 58530, A
71	19	95.0	89	12	US-10-282-122A-66960	Sequence 66960, A
72	19	95.0	89	12	US-10-282-122A-68974	Sequence 68774, A
73	19	95.0	89	12	US-10-424-599-180307	Sequence 180307, A
74	19	95.0	89	14	US-10-097-111-360	Sequence 360, App
75	19	95.0	90	12	US-10-424-599-245830	Sequence 245830, A
76	19	95.0	93	12	US-10-424-599-300533	Sequence 300533, A
77	19	95.0	93	12	US-10-424-599-204269	Sequence 204269, A
78	19	95.0	94	15	US-10-372-003A-79	Sequence 79, Appl
79	19	95.0	99	15	US-10-264-049-2775	Sequence 2775, Ap
80	19	95.0	100	12	US-10-338-977-7495	Sequence 7495, Ap
81	19	95.0	101	12	US-10-335-977-7494	Sequence 7494, Ap
82	19	95.0	103	12	US-10-424-599-196121	Sequence 196121, A
83	19	95.0	104	12	US-10-424-599-245933	Sequence 245933, A
84	19	95.0	104	16	US-10-294-445-30	Sequence 30, Appl
85	19	95.0	105	14	US-10-291-851-90	Sequence 90, Appl
86	19	95.0	107	9	US-09-738-626-4520	Sequence 4520, Ap
87	19	95.0	109	12	US-10-363-616-367	Sequence 367, App
88	19	95.0	109	15	US-10-264-237-1538	Sequence 1538, Ap



89	19	95.0	110	12	US-10-424-599-144932	Sequence 144932,	162	19	95.0	189	12	US-10-424-599-209651	Sequence 209651,
90	19	95.0	110	12	US-10-424-599-257728	Sequence 257728,	163	19	95.0	189	14	US-10-238-075-4037	Sequence 1037, Ap
91	19	95.0	111	12	US-10-424-599-192406	Sequence 192406,	164	19	95.0	191	12	US-10-282-122A-77681	Sequence 77681, A
92	19	95.0	112	12	US-10-424-599-189066	Sequence 189066,	165	19	95.0	192	12	US-10-424-599-269537	Sequence 269537,
93	19	95.0	112	12	US-10-424-599-232263	Sequence 232263,	166	19	95.0	193	9	US-09-888-911-4	Sequence 4, Appli
94	19	95.0	117	12	US-10-424-599-206883	Sequence 206883,	167	19	95.0	193	12	US-10-425-114-43750	Sequence 43750, A
95	19	95.0	117	12	US-10-424-599-221584	Sequence 221584,	168	19	95.0	193	14	US-10-156-761-13355	Sequence 13355, A
96	19	95.0	118	12	US-10-424-599-271082	Sequence 271082,	169	19	95.0	193	15	US-10-377-072-38	Sequence 38, Appl
97	19	95.0	119	12	US-10-424-599-240644	Sequence 240644,	170	19	95.0	195	12	US-10-424-599-227828	Sequence 227828,
98	19	95.0	119	14	US-10-263-568-12	Sequence 12, Appl	171	19	95.0	195	12	US-10-424-599-231551	Sequence 231551,
99	19	95.0	120	12	US-10-424-599-189718	Sequence 189718,	172	19	95.0	195	12	US-10-424-599-238939	Sequence 238939,
100	19	95.0	120	12	US-10-424-599-276785	Sequence 276785,	173	19	95.0	195	12	US-10-369-493-3234	Sequence 3234, Ap
101	19	95.0	122	11	US-09-864-408A-526	Sequence 526, App	174	19	95.0	198	9	US-09-966-880A-2	Sequence 2, Appli
102	19	95.0	122	12	US-10-424-599-173522	Sequence 173522,	175	19	95.0	198	12	US-10-425-114-47425	Sequence 47425, A
103	19	95.0	127	12	US-10-424-599-178767	Sequence 178767,	176	19	95.0	199	12	US-10-403-877-85	Sequence 85, Appl
104	19	95.0	129	16	US-10-399-636-68	Sequence 68, Appl	177	19	95.0	200	12	US-10-424-599-176742	Sequence 176742,
105	19	95.0	130	12	US-10-424-599-184879	Sequence 184879,	178	19	95.0	200	12	US-10-424-599-220990	Sequence 220990,
106	19	95.0	130	12	US-10-424-599-199911	Sequence 199911,	179	19	95.0	201	12	US-10-276-162-2	Sequence 2, Appli
107	19	95.0	131	9	US-09-738-626-5204	Sequence 5204, Ap	180	19	95.0	201	15	US-10-085-198-134	Sequence 134, App
108	19	95.0	132	12	US-10-243-559-932	Sequence 932, App	181	19	95.0	202	12	US-10-283-122A-60863	Sequence 60863 A
109	19	95.0	133	9	US-09-864-761-33708	Sequence 33708, A	182	19	95.0	204	12	US-10-424-599-231852	Sequence 231852,
110	19	95.0	133	12	US-10-424-599-158331	Sequence 158331,	183	19	95.0	204	12	US-10-424-599-265941	Sequence 265941,
111	19	95.0	134	9	US-09-815-342-10513	Sequence 10513, A	184	19	95.0	204	12	US-10-634-548-29	Sequence 29, Appl
112	19	95.0	135	12	US-10-424-599-271153	Sequence 271153, A	185	19	95.0	204	12	US-10-634-548-31	Sequence 31, Appl
113	19	95.0	135	12	US-10-425-114-53378	Sequence 53378, A	186	19	95.0	205	12	US-09-922-217-11308	Sequence 1108, Ap
114	19	95.0	138	12	US-10-112-944-912	Sequence 912, App	187	19	95.0	205	12	US-10-296-115-1330	Sequence 1330, Ap
115	19	95.0	139	14	US-10-156-761-8842	Sequence 8842, Ap	188	19	95.0	205	12	US-10-335-977-7448	Sequence 7448, Ap
116	19	95.0	139	15	US-10-385-415-120	Sequence 120, App	189	19	95.0	205	12	US-10-335-977-7980	Sequence 7980, Ap
117	19	95.0	140	15	US-10-108-260A-3390	Sequence 3390, Ap	190	19	95.0	205	13	US-10-023-380-1108	Sequence 1108, Ap
118	19	95.0	143	15	US-10-385-415-119	Sequence 119, App	191	19	95.0	208	12	US-10-023-380-1108	Sequence 4800, A
119	19	95.0	143	15	US-10-013-379-11	Sequence 11, Appl	192	19	95.0	208	14	US-10-282-122A-44800	Sequence 13562, A
120	19	95.0	145	15	US-10-289-762-167	Sequence 167, App	193	19	95.0	209	12	US-10-283-122A-72561	Sequence 72561, A
121	19	95.0	146	12	US-10-282-122A-70260	Sequence 70260, A	194	19	95.0	209	12	US-10-424-599-367539	Sequence 367539,
122	19	95.0	146	15	US-10-139-79A-87	Sequence 87, Appl	195	19	95.0	212	14	US-10-127-032-130	Sequence 130, App
123	19	95.0	148	12	US-10-424-599-143742	Sequence 143742,	196	19	95.0	213	9	US-09-943-671-17	Sequence 17, Appl
124	19	95.0	148	12	US-10-424-599-219032	Sequence 219032,	197	19	95.0	213	9	US-09-738-626-3673	Sequence 3673, Ap
125	19	95.0	152	12	US-10-424-599-283972	Sequence 283972,	198	19	95.0	213	14	US-10-043-487-364	Sequence 364, App
126	19	95.0	152	13	US-10-016-157A-151	Sequence 151, App	199	19	95.0	213	15	US-10-405-877-17	Sequence 17, Appl
127	19	95.0	156	12	US-10-424-599-254161	Sequence 254161,	200	19	95.0	214	14	US-10-106-698-5887	Sequence 5887, Ap
128	19	95.0	157	14	US-10-238-075-357	Sequence 357, App	201	19	95.0	215	14	US-09-738-626-5708	Sequence 5708, Ap
129	19	95.0	159	12	US-10-424-599-256669	Sequence 256669,	202	19	95.0	215	12	US-10-283-122A-62976	Sequence 62976, A
130	19	95.0	163	12	US-10-424-599-169320	Sequence 169320,	203	19	95.0	215	14	US-10-323-671-69	Sequence 69, Appl
131	19	95.0	164	9	US-09-738-973-203	Sequence 203, App	204	19	95.0	217	12	US-10-243-542-890	Sequence 890, App
132	19	95.0	164	11	US-09-854-133-203	Sequence 203, App	205	19	95.0	217	12	US-10-283-122A-68326	Sequence 68326, A
133	19	95.0	164	13	US-09-864-408A-1148	Sequence 1148, Ap	206	19	95.0	218	12	US-10-282-122A-78537	Sequence 78537, A
134	19	95.0	164	14	US-10-144-649A-203	Sequence 203, App	207	19	95.0	219	12	US-10-424-599-154224	Sequence 154224,
135	19	95.0	165	12	US-10-424-599-148999	Sequence 148999,	208	19	95.0	219	14	US-10-156-761-9995	Sequence 9995, Ap
136	19	95.0	165	12	US-10-424-599-234773	Sequence 234773,	209	19	95.0	220	12	US-10-424-599-162041	Sequence 162041,
137	19	95.0	167	9	US-09-925-300-981	Sequence 981, App	210	19	95.0	220	12	US-10-384-080-76	Sequence 76, Appl
138	19	95.0	168	12	US-10-424-599-198953	Sequence 198953,	211	19	95.0	221	12	US-10-424-599-266598	Sequence 266598,
139	19	95.0	172	12	US-10-425-114-57575	Sequence 57575, A	212	19	95.0	221	14	US-10-156-761-12750	Sequence 12750, Ap
140	19	95.0	172	14	US-10-101-489-16	Sequence 16, Appl	213	19	95.0	222	9	US-09-925-300-1639	Sequence 1639, Ap
141	19	95.0	173	12	US-10-424-599-158029	Sequence 158029,	214	19	95.0	222	12	US-10-424-599-157122	Sequence 157122,
142	19	95.0	175	12	US-10-335-977-5868	Sequence 5868, Ap	215	19	95.0	222	12	US-10-425-114-36974	Sequence 36974, A
143	19	95.0	175	14	US-10-235-994-12	Sequence 12, Appl	216	19	95.0	222	12	US-10-425-114-52549	Sequence 52549, A
144	19	95.0	175	14	US-10-156-761-13575	Sequence 12575, A	217	19	95.0	224	9	US-09-738-626-4621	Sequence 4621, Ap
145	19	95.0	175	14	US-10-424-599-184	Sequence 184, App	218	19	95.0	224	12	US-10-424-599-151189	Sequence 151189,
146	19	95.0	176	9	US-09-864-761-33919	Sequence 33919, App	219	19	95.0	224	15	US-10-094-749-1944	Sequence 1944, Ap
147	19	95.0	178	9	US-09-925-301-914	Sequence 925, App	220	19	95.0	225	12	US-10-424-599-262650	Sequence 262650,
148	19	95.0	178	15	US-10-364-049-2578	Sequence 2578, Ap	221	19	95.0	225	14	US-10-351-334-155	Sequence 155, App
149	19	95.0	180	12	US-10-424-599-170244	Sequence 170244,	222	19	95.0	225	12	US-10-156-761-14916	Sequence 14916 A
150	19	95.0	180	14	US-10-156-761-13228	Sequence 12228, A	223	19	95.0	225	14	US-10-369-493-11674	Sequence 11674, A
151	19	95.0	180	14	US-10-335-977-7975	Sequence 7975, App	224	19	95.0	225	15	US-10-369-493-14289	Sequence 14289, A
152	19	95.0	182	12	US-10-424-599-156-17	Sequence 17, Appl	225	19	95.0	225	15	US-10-369-493-14859	Sequence 14859, A
153	19	95.0	187	9	US-09-804-156-17	Sequence 9, Appli	226	19	95.0	226	9	US-09-796-149-25	Sequence 5, Appli
154	19	95.0	187	9	US-09-946-633-9	Sequence 9, Appli	227	19	95.0	226	12	US-10-424-599-239616	Sequence 239616,
155	19	95.0	187	13	US-10-125-459-9	Sequence 9, Appli	228	19	95.0	227	16	US-10-275-505-7	Sequence 7, Appli
156	19	95.0	187	14	US-10-067-76-17	Sequence 17, Appl	229	19	95.0	228	12	US-10-425-114-47426	Sequence 47426, A
157	19	95.0	187	14	US-10-319-519-17	Sequence 17, Appl	230	19	95.0	228	13	US-10-051-325-8	Sequence 8, Appli
158	19	95.0	188	9	US-09-734-569-134	Sequence 134, App	231	19	95.0	230	9	US-09-738-626-4087	Sequence 143683,
159	19	95.0	188	12	US-10-282-122A-76586	Sequence 76586, A	232	19	95.0	230	9	US-10-424-599-143683	Sequence 4027, Ap
160	19	95.0	188	12	US-10-282-122A-78177	Sequence 78177, A	233	19	95.0	231	12	US-10-425-114-53174	Sequence 53174, A
161	19	95.0	188	12	US-10-424-599-256741	Sequence 256741,	234	19	95.0	231	12	US-10-425-114-57318	Sequence 57318, A

235	19	95.0	232	10	US-09-252-088-40	Sequence 40, Appl	308	19	95.0	270	9	US-09-738-626-3943	Sequence 3943, Ap
236	19	95.0	232	15	US-10-340-792-40	Sequence 40, Appl	309	19	95.0	270	12	US-10-262-511-148	Sequence 148, App
237	19	95.0	234	12	US-10-425-114-45098	Sequence 45098, A	310	19	95.0	270	15	US-10-269-493-10566	Sequence 10566, A
238	19	95.0	234	12	US-10-425-114-63330	Sequence 63330, A	311	19	95.0	271	15	US-10-264-237-2658	Sequence 2658, Ap
239	19	95.0	236	12	US-10-425-114-64490	Sequence 64490, A	312	19	95.0	272	10	US-09-764-891-2960	Sequence 2960, Ap
240	19	95.0	236	9	US-09-815-242-11738	Sequence 11738, A	313	19	95.0	272	12	US-10-282-122A-70355	Sequence 70355, A
241	19	95.0	236	12	US-10-412-698B-1410	Sequence 1410, Ap	314	19	95.0	273	14	US-10-182-447-3	Sequence 3, Appl1
242	19	95.0	236	12	US-10-421-138A-308	Sequence 308, App	315	19	95.0	274	12	US-10-282-122A-51559	Sequence 51559, A
243	19	95.0	236	15	US-10-374-780A-1221	Sequence 1221, Ap	316	19	95.0	274	12	US-10-282-122A-74371	Sequence 74371, A
244	19	95.0	238	12	US-10-282-122A-70662	Sequence 70662, A	317	19	95.0	275	12	US-10-424-599-146071	Sequence 52045, A
245	19	95.0	238	12	US-10-282-122A-72399	Sequence 72399, A	318	19	95.0	275	12	US-10-282-122A-65461	Sequence 65461, A
246	19	95.0	238	12	US-10-137-036-68	Sequence 68, Appl	319	19	95.0	278	12	US-10-282-122A-65761	Sequence 65761, A
247	19	95.0	239	12	US-10-424-599-283468	Sequence 283468, A	320	19	95.0	278	12	US-10-282-122A-65761	Sequence 11888, A
248	19	95.0	240	14	US-10-338-777-39	Sequence 39, Appl	321	19	95.0	279	15	US-10-369-493-11888	Sequence 4505, Ap
249	19	95.0	241	12	US-10-282-122A-56234	Sequence 56234, A	322	19	95.0	280	9	US-09-738-626-4503	Sequence 20769, A
250	19	95.0	241	12	US-10-425-114-36715	Sequence 36715, A	323	19	95.0	280	15	US-10-369-493-20769	Sequence 58, Appl
251	19	95.0	242	12	US-10-425-114-47756	Sequence 47756, A	324	19	95.0	281	13	US-10-059-964-58	Sequence 58, Appl
252	19	95.0	242	12	US-10-425-114-44756	Sequence 44, Appl	325	19	95.0	281	14	US-10-314-639-58	Sequence 9102, Ap
253	19	95.0	243	10	US-09-870-406A-44	Sequence 257488, A	326	19	95.0	281	14	US-10-156-761-9102	Sequence 1694, Ap
254	19	95.0	243	12	US-10-424-599-257488	Sequence 257488, A	327	19	95.0	282	9	US-09-925-300-1694	Sequence 47275, A
255	19	95.0	243	14	US-10-159-901-44	Sequence 44, Appl	328	19	95.0	282	12	US-10-282-122A-47275	Sequence 12355, A
256	19	95.0	244	12	US-10-282-122A-72187	Sequence 72187, A	329	19	95.0	282	14	US-10-156-761-12355	Sequence 5203, Ap
257	19	95.0	244	12	US-10-424-599-158126	Sequence 158126, A	330	19	95.0	283	8	US-08-781-986A-5203	Sequence 43988, A
258	19	95.0	244	12	US-10-424-599-271081	Sequence 271081, A	331	19	95.0	283	12	US-10-425-114-43988	Sequence 5203, Ap
259	19	95.0	246	9	US-09-771-161A-129	Sequence 129, App	332	19	95.0	283	12	US-10-329-624-5203	Sequence 58685, A
260	19	95.0	246	15	US-10-156-761-13663	Sequence 13663, A	333	19	95.0	284	12	US-10-425-114-58685	Sequence 49530, A
261	19	95.0	246	15	US-10-421-138A-311	Sequence 311, App	334	19	95.0	285	12	US-10-282-122A-49530	Sequence 69922, A
262	19	95.0	246	15	US-10-374-780A-1244	Sequence 1244, Ap	335	19	95.0	285	12	US-10-282-122A-69922	Sequence 64235, A
263	19	95.0	248	12	US-10-425-114-44065	Sequence 44065, A	336	19	95.0	285	12	US-10-425-114-64235	Sequence 69316, A
264	19	95.0	249	9	US-09-832-312-9	Sequence 9, Appl1	337	19	95.0	286	12	US-10-282-122A-69316	Sequence 144, App
265	19	95.0	249	11	US-09-829-495-9	Sequence 9, Appl1	338	19	95.0	286	12	US-10-262-511-144	Sequence 147161, A
266	19	95.0	249	14	US-10-156-761-9204	Sequence 9204, Ap	339	19	95.0	287	12	US-10-424-599-147161	Sequence 20884, A
267	19	95.0	249	15	US-10-409-701-15	Sequence 15, Appl	340	19	95.0	287	15	US-10-369-493-20884	Sequence 66912, A
268	19	95.0	249	15	US-10-310-154-606	Sequence 606, App	341	19	95.0	288	12	US-10-282-122A-66912	Sequence 66363, A
269	19	95.0	253	9	US-09-815-242-10279	Sequence 10279, A	342	19	95.0	288	9	US-09-815-242-11193	Sequence 11193, A
270	19	95.0	253	15	US-10-369-493-23537	Sequence 23537, A	343	19	95.0	289	12	US-10-282-122A-58429	Sequence 58429, A
271	19	95.0	254	12	US-10-424-599-158554	Sequence 158554, A	344	19	95.0	289	12	US-10-425-114-47196	Sequence 47196, A
272	19	95.0	254	12	US-10-424-599-189853	Sequence 189853, A	345	19	95.0	290	12	US-10-424-599-196081	Sequence 196081, A
273	19	95.0	254	12	US-10-262-511-150	Sequence 150, App	346	19	95.0	290	12	US-10-262-511-146	Sequence 146, App
274	19	95.0	254	15	US-10-369-493-11429	Sequence 11429, Ap	347	19	95.0	290	15	US-10-369-493-15865	Sequence 15865, A
275	19	95.0	255	13	US-10-086-623-40	Sequence 40, Appl	348	19	95.0	290	15	US-10-041-018-27	Sequence 27, Appl
276	19	95.0	255	14	US-10-260-539-40	Sequence 40, Appl	349	19	95.0	291	12	US-10-017-161-2430	Sequence 2430, Ap
277	19	95.0	256	15	US-10-369-493-19054	Sequence 19054, A	350	19	95.0	291	14	US-10-092-798-2070	Sequence 2070, Ap
278	19	95.0	256	15	US-10-421-138A-312	Sequence 312, App	351	19	95.0	291	15	US-10-264-237-2770	Sequence 5341, Ap
279	19	95.0	256	15	US-10-374-780A-1247	Sequence 1247, Ap	352	19	95.0	292	15	US-10-369-493-5341	Sequence 15494, A
280	19	95.0	257	9	US-09-738-626-6642	Sequence 6642, Ap	353	19	95.0	292	15	US-10-282-122A-76466	Sequence 76466, A
281	19	95.0	257	14	US-10-106-698-5221	Sequence 5221, Ap	354	19	95.0	293	12	US-10-156-761-14726	Sequence 14726, A
282	19	95.0	257	15	US-10-139-794-16	Sequence 16, Appl	355	19	95.0	293	14	US-10-299-867-20	Sequence 12, Appl
283	19	95.0	258	9	US-09-815-242-11889	Sequence 11889, A	356	19	95.0	294	14	US-10-236-699-12	Sequence 155, Appl
284	19	95.0	258	12	US-10-282-122A-66426	Sequence 66426, A	357	19	95.0	294	14	US-10-306-762-155	Sequence 13821, A
285	19	95.0	258	12	US-10-424-599-248969	Sequence 248969, A	358	19	95.0	294	15	US-10-369-493-13821	Sequence 141, App
286	19	95.0	258	14	US-10-097-111-303	Sequence 303, App	359	19	95.0	295	14	US-10-306-762-141	Sequence 50243, A
287	19	95.0	260	12	US-10-425-114-46566	Sequence 46566, A	360	19	95.0	295	14	US-10-282-122A-50243	Sequence 60194, A
288	19	95.0	261	10	US-09-893-519A-19	Sequence 19, Appl	361	19	95.0	297	12	US-10-282-122A-60194	Sequence 41342, A
289	19	95.0	261	12	US-10-282-122A-69377	Sequence 69377, A	362	19	95.0	297	12	US-10-425-114-41342	Sequence 35, Appl
290	19	95.0	261	14	US-10-011-364-10	Sequence 10, Appl	363	19	95.0	299	9	US-09-393-634-35	Sequence 1, Appl1
291	19	95.0	261	15	US-10-369-493-12238	Sequence 12238, A	364	19	95.0	299	10	US-09-510-332-1	Sequence 73, Appl
292	19	95.0	261	15	US-10-321-962-10	Sequence 10, Appl	365	19	95.0	299	10	US-09-812-522-73	Sequence 33, Appl
293	19	95.0	263	9	US-10-424-599-254101	Sequence 254101, A	366	19	95.0	299	13	US-10-364-861-35	Sequence 46, Appl
294	19	95.0	263	10	US-09-800-729-88	Sequence 88, Appl	367	19	95.0	299	13	US-10-059-964-46	Sequence 46, Appl
295	19	95.0	263	11	US-09-746-783-2	Sequence 2, Appl1	368	19	95.0	299	13	US-10-314-639-46	Sequence 35, Appl
296	19	95.0	263	14	US-10-182-447-1	Sequence 1, Appl1	369	19	95.0	299	14	US-10-383-982-35	Sequence 3734, Ap
297	19	95.0	264	12	US-10-282-122A-52414	Sequence 52414, A	370	19	95.0	303	9	US-09-738-626-3734	Sequence 53342, A
298	19	95.0	265	12	US-10-425-114-41437	Sequence 41437, A	371	19	95.0	303	12	US-10-425-114-53342	Sequence 219921, A
299	19	95.0	265	12	US-10-425-114-62005	Sequence 62005, A	372	19	95.0	304	12	US-10-425-114-62005	Sequence 62081, A
300	19	95.0	265	12	US-10-424-599-352581	Sequence 352581, A	373	19	95.0	304	15	US-10-369-493-9667	Sequence 9667, Ap
301	19	95.0	266	12	US-10-282-122A-78178	Sequence 78178, A	374	19	95.0	304	15	US-10-425-114-62081	Sequence 6591, Ap
302	19	95.0	267	12	US-10-282-122A-78178	Sequence 45104, A	375	19	95.0	305	9	US-09-738-626-6591	Sequence 62473, A
303	19	95.0	267	12	US-10-425-114-45104	Sequence 45104, A	376	19	95.0	306	12	US-10-425-114-62473	Sequence 64440, A
304	19	95.0	267	12	US-10-425-114-46738	Sequence 46738, A	377	19	95.0	306	12	US-10-425-114-64440	Sequence 67258, A
305	19	95.0	267	15	US-10-421-138A-315	Sequence 315, App	378	19	95.0	306	12	US-10-425-114-64440	Sequence 67258, A
306	19	95.0	268	16	US-10-258-662-19	Sequence 16, Appl	379	19	95.0	306	12	US-10-425-114-64440	Sequence 67258, A
307	19	95.0	269	15	US-10-104-047-2182	Sequence 2182, Ap	380	19	95.0	306	12	US-10-425-114-67258	Sequence 67258, A

381	19	95.0	306	14	US-10-091-007-40	Sequence 40, Appl	454	19	95.0	321	14	US-10-025-806-32	Sequence 32, Appl
382	19	95.0	306	15	US-10-320-797-3084	Sequence 3084, Ap	455	19	95.0	322	12	US-10-283-132A-53145	Sequence 53145, A
383	19	95.0	307	14	US-10-156-761-8528	Sequence 8528, Ap	456	19	95.0	322	14	US-10-270-861-7	Sequence 7, Appl
384	19	95.0	308	9	US-09-925-299-930	Sequence 930, App	457	19	95.0	322	14	US-10-270-861-11	Sequence 11, Appl
385	19	95.0	308	9	US-09-738-626-6728	Sequence 6728, Ap	458	19	95.0	323	12	US-10-282-122A-69906	Sequence 69906, A
386	19	95.0	308	14	US-09-525-299-930	Sequence 930, App	459	19	95.0	323	12	US-10-424-599-238938	Sequence 238938, A
387	19	95.0	308	14	US-10-032-201B-253	Sequence 253, App	460	19	95.0	324	12	US-10-282-122A-57182	Sequence 57182, A
388	19	95.0	309	12	US-10-425-114-45522	Sequence 45522, A	461	19	95.0	324	12	US-10-282-122A-61308	Sequence 61308, A
389	19	95.0	309	14	US-10-023-601-96	Sequence 96, Appl	462	19	95.0	324	14	US-10-156-761-13399	Sequence 13399, A
390	19	95.0	309	15	US-10-369-493-7640	Sequence 7640, Ap	463	19	95.0	324	14	US-10-156-761-14083	Sequence 14083, A
391	19	95.0	309	15	US-10-369-493-11118	Sequence 11118, A	464	19	95.0	325	12	US-10-262-511-152	Sequence 152, App
392	19	95.0	310	9	US-09-826-309-2	Sequence 2, Appli	465	19	95.0	325	14	US-10-270-861-1	Sequence 1, Appl
393	19	95.0	310	12	US-10-282-122A-67405	Sequence 67405, A	466	19	95.0	325	14	US-10-270-861-13	Sequence 13, Appl
394	19	95.0	310	12	US-10-424-599-163497	Sequence 163497, A	467	19	95.0	326	15	US-10-369-493-1001	Sequence 1001, Ap
395	19	95.0	310	12	US-10-424-599-230916	Sequence 230916, A	468	19	95.0	327	12	US-10-424-599-214738	Sequence 214738, A
396	19	95.0	310	15	US-10-191-121-18	Sequence 18, Appl	469	19	95.0	327	12	US-10-425-114-52852	Sequence 52852, A
397	19	95.0	310	16	US-10-360-534-16	Sequence 16, Appl	470	19	95.0	327	14	US-10-314-657-8	Sequence 8, Appli
398	19	95.0	311	9	US-09-896-055-9	Sequence 9, Appli	471	19	95.0	328	9	US-09-815-242-10935	Sequence 10935, A
399	19	95.0	311	9	US-09-738-626-3671	Sequence 3671, Ap	472	19	95.0	328	10	US-09-765-061B-75	Sequence 75, Appl
400	19	95.0	311	10	US-09-804-291-9	Sequence 9, Appli	473	19	95.0	328	10	US-09-765-061B-76	Sequence 76, Appl
401	19	95.0	311	10	US-09-832-522-20	Sequence 20, Appl	474	19	95.0	328	10	US-09-765-061B-76	Sequence 76, Appl
402	19	95.0	311	10	US-09-795-271-25	Sequence 25, Appl	475	19	95.0	328	12	US-10-282-122A-56878	Sequence 56878, A
403	19	95.0	311	11	US-09-981-151A-75	Sequence 75, Appl	476	19	95.0	328	15	US-10-369-493-20838	Sequence 20838, A
404	19	95.0	311	11	US-09-844-861A-12	Sequence 12, Appl	477	19	95.0	330	9	US-09-771-161A-128	Sequence 128, App
405	19	95.0	311	12	US-10-343-650A-364	Sequence 260, App	478	19	95.0	331	9	US-09-815-242-11761	Sequence 11761, A
406	19	95.0	311	12	US-10-343-650A-364	Sequence 364, App	479	19	95.0	331	12	US-10-424-599-176898	Sequence 176898, A
407	19	95.0	311	14	US-10-023-597-52	Sequence 52, Appl	480	19	95.0	331	12	US-10-425-114-43781	Sequence 43781, A
408	19	95.0	311	14	US-10-017-161-280	Sequence 280, App	481	19	95.0	334	15	US-10-282-122A-71145	Sequence 71145, A
409	19	95.0	311	14	US-10-017-161-284	Sequence 284, App	482	19	95.0	335	15	US-10-264-049-2446	Sequence 2446, Ap
410	19	95.0	311	15	US-10-005-041A-20	Sequence 20, Appl	483	19	95.0	335	12	US-09-984-271-118	Sequence 118, App
411	19	95.0	311	15	US-10-292-798-250	Sequence 250, App	484	19	95.0	335	12	US-10-425-114-39152	Sequence 39152, A
412	19	95.0	311	15	US-10-292-798-254	Sequence 254, App	485	19	95.0	335	15	US-09-984-271-118	Sequence 118, App
413	19	95.0	312	9	US-09-866-055-187	Sequence 187, App	486	19	95.0	335	15	US-10-369-493-18792	Sequence 18792, A
414	19	95.0	312	9	US-09-738-626-5007	Sequence 5007, Ap	487	19	95.0	336	12	US-10-282-122A-67767	Sequence 67767, A
415	19	95.0	312	10	US-09-804-291-187	Sequence 187, App	488	19	95.0	336	12	US-10-282-122A-78253	Sequence 78253, A
416	19	95.0	312	12	US-10-282-122A-53997	Sequence 53997, A	489	19	95.0	336	15	US-10-425-114-46555	Sequence 46555, A
417	19	95.0	312	12	US-10-182-822A-17	Sequence 17, Appl	490	19	95.0	337	10	US-09-793-705-42	Sequence 42, Appl
418	19	95.0	312	12	US-10-331-079-34	Sequence 34, Appl	491	19	95.0	337	12	US-10-425-114-57499	Sequence 57499, A
419	19	95.0	312	12	US-10-343-650A-656	Sequence 556, App	492	19	95.0	337	12	US-10-425-114-59818	Sequence 59818, A
420	19	95.0	312	14	US-10-017-161-50	Sequence 50, App	493	19	95.0	337	12	US-10-425-114-63730	Sequence 63730, A
421	19	95.0	312	14	US-10-017-161-50	Sequence 50, App	494	19	95.0	337	15	US-10-369-493-11195	Sequence 11195, A
422	19	95.0	312	15	US-10-387-629-16	Sequence 16, Appl	495	19	95.0	338	9	US-09-852-976-1	Sequence 1, Appli
423	19	95.0	312	15	US-10-292-798-42	Sequence 42, Appl	496	19	95.0	338	12	US-10-425-114-53202	Sequence 53202, A
424	19	95.0	313	15	US-10-369-493-11089	Sequence 11089, A	497	19	95.0	338	12	US-10-425-114-61153	Sequence 61153, A
425	19	95.0	314	12	US-10-282-122A-45385	Sequence 45385, A	498	19	95.0	338	12	US-10-425-114-62750	Sequence 62750, A
426	19	95.0	314	12	US-10-424-599-181578	Sequence 181578, A	499	19	95.0	338	12	US-10-425-114-66618	Sequence 66618, A
427	19	95.0	314	12	US-10-424-599-185947	Sequence 185947, A	500	19	95.0	339	9	US-09-832-312-3	Sequence 3, Appli
428	19	95.0	315	12	US-10-424-599-185947	Sequence 185947, A	501	19	95.0	339	9	US-09-832-312-34	Sequence 34, Appl
429	19	95.0	315	12	US-10-023-597-102	Sequence 102, App	502	19	95.0	339	9	US-09-832-312-36	Sequence 36, Appl
430	19	95.0	315	14	US-10-238-075-812	Sequence 812, App	503	19	95.0	339	9	US-09-832-312-38	Sequence 38, Appl
431	19	95.0	315	14	US-10-311-626-1	Sequence 1, Appli	504	19	95.0	339	9	US-09-832-312-40	Sequence 40, Appl
432	19	95.0	315	16	US-10-309-290-138	Sequence 138, App	505	19	95.0	339	11	US-09-829-495-3	Sequence 3, Appli
433	19	95.0	316	12	US-10-072-012-407	Sequence 407, App	506	19	95.0	339	11	US-09-829-495-34	Sequence 34, Appl
434	19	95.0	316	14	US-10-326-699-16	Sequence 16, Appl	507	19	95.0	339	11	US-09-829-495-35	Sequence 35, Appl
435	19	95.0	317	12	US-10-425-114-43447	Sequence 43447, A	508	19	95.0	339	11	US-09-829-495-38	Sequence 38, Appl
436	19	95.0	318	10	US-09-795-271-40	Sequence 40, Appl	509	19	95.0	339	11	US-09-829-495-38	Sequence 40, Appl
437	19	95.0	318	10	US-09-907-218-63	Sequence 63, Appl	510	19	95.0	339	14	US-10-157-031-387	Sequence 387, App
438	19	95.0	318	11	US-09-965-422-41	Sequence 41, Appl	511	19	95.0	339	14	US-10-156-761-14680	Sequence 14680, A
439	19	95.0	318	11	US-09-981-566A-42	Sequence 42, Appl	512	19	95.0	339	14	US-10-053-248-34	Sequence 34, Appl
440	19	95.0	318	11	US-09-981-566A-47	Sequence 47, Appl	513	19	95.0	339	14	US-10-446-826-5	Sequence 5, Appli
441	19	95.0	318	11	US-09-981-566A-47	Sequence 47, Appl	514	19	95.0	342	12	US-10-425-114-48872	Sequence 48872, A
442	19	95.0	318	14	US-10-032-201B-284	Sequence 284, App	515	19	95.0	342	12	US-10-276-774-2289	Sequence 2289, Ap
443	19	95.0	318	14	US-10-375-611-1	Sequence 1, Appli	516	19	95.0	342	12	US-09-805-020-49	Sequence 49, Appl
444	19	95.0	319	9	US-09-832-312-5	Sequence 5, Appli	517	19	95.0	342	14	US-10-300-075-158	Sequence 158, Ap
445	19	95.0	319	11	US-09-829-495-5	Sequence 5, Appli	518	19	95.0	345	9	US-09-525-300-1188	Sequence 1188, Ap
446	19	95.0	319	12	US-10-425-114-70828	Sequence 70828, A	519	19	95.0	345	12	US-10-424-599-223433	Sequence 223433, A
447	19	95.0	319	13	US-10-001-426-4	Sequence 4, Appli	520	19	95.0	346	10	US-09-953-348-107	Sequence 107, App
448	19	95.0	319	13	US-10-003-014-4	Sequence 4, Appli	521	19	95.0	346	12	US-10-425-114-66218	Sequence 66218, A
449	19	95.0	319	14	US-10-236-699-10	Sequence 10, Appl	522	19	95.0	346	12	US-10-425-114-67904	Sequence 67904, A
450	19	95.0	320	12	US-10-424-599-186648	Sequence 186648, A	523	19	95.0	346	14	US-10-267-255-107	Sequence 107, App
451	19	95.0	320	12	US-10-335-977-5813	Sequence 5813, Ap	524	19	95.0	347	9	US-09-738-626-6922	Sequence 6922, Ap
452	19	95.0	320	16	US-10-461-990-28	Sequence 28, Appl	525	19	95.0	347	12	US-10-425-114-52822	Sequence 52822, A
453	19	95.0	321	12	US-10-424-599-163498	Sequence 163498, A	526	19	95.0	347	14	US-10-306-762-240	Sequence 240, App

527	19	95.0	347	15	US-10-264-043-2882	Sequence 2882, Ap	600	19	95.0	384	15	US-10-369-493-18459	Sequence 18459, A
528	19	95.0	349	12	US-10-335-977-5814	Sequence 5814, Ap	601	19	95.0	385	12	US-10-425-114-67317	Sequence 67317, A
529	19	95.0	349	16	US-10-389-566-1121	Sequence 1121, Ap	602	19	95.0	385	15	US-10-369-493-6719	Sequence 6719, A
530	19	95.0	349	16	US-10-389-566-2148	Sequence 2148, Ap	603	19	95.0	386	9	US-09-815-242-11041	Sequence 11041, A
531	19	95.0	351	12	US-10-282-122A-46888	Sequence 46888, A	604	19	95.0	386	12	US-10-282-122A-58201	Sequence 58201, A
532	19	95.0	351	15	US-10-369-493-10659	Sequence 10659, A	605	19	95.0	386	15	US-10-369-493-4047	Sequence 4047, Ap
533	19	95.0	351	15	US-10-369-493-10659	Sequence 10659, A	606	19	95.0	389	12	US-10-425-114-59654	Sequence 59654, A
534	19	95.0	352	12	US-10-282-122A-60673	Sequence 60673, A	607	19	95.0	389	12	US-10-166-225A-82	Sequence 82, Appl
535	19	95.0	352	15	US-10-369-493-20240	Sequence 20240, A	608	19	95.0	389	14	US-10-166-225A-83	Sequence 83, Appl
536	19	95.0	352	16	US-10-389-566-2165	Sequence 2165, Ap	609	19	95.0	389	14	US-10-166-225A-83	Sequence 83, Appl
537	19	95.0	353	14	US-10-270-861-9	Sequence 9, Appl	610	19	95.0	389	15	US-10-369-493-20787	Sequence 20787, A
538	19	95.0	354	16	US-10-616-624-4	Sequence 4, Appl	611	19	95.0	390	15	US-10-369-493-8879	Sequence 8879, Ap
539	19	95.0	355	14	US-10-106-698-4846	Sequence 4846, Ap	612	19	95.0	390	15	US-10-369-493-18574	Sequence 18574, A
540	19	95.0	357	12	US-10-363-618-438	Sequence 438, App	613	19	95.0	391	15	US-10-369-493-3160	Sequence 3160, Ap
541	19	95.0	358	12	US-10-087-192-1845	Sequence 1845, Ap	614	19	95.0	391	15	US-10-369-493-18574	Sequence 18574, A
542	19	95.0	359	14	US-10-032-585-7893	Sequence 7893, Ap	615	19	95.0	392	10	US-08-785-0618-77	Sequence 77, Appl
543	19	95.0	360	12	US-10-425-114-65637	Sequence 65637, A	616	19	95.0	392	12	US-10-369-493-11600	Sequence 11600, A
544	19	95.0	363	12	US-10-425-114-61874	Sequence 61874, A	617	19	95.0	392	15	US-10-369-493-14301	Sequence 14301, A
545	19	95.0	364	12	US-10-424-599-214841	Sequence 214841, A	618	19	95.0	392	15	US-10-369-493-14467	Sequence 14467, A
546	19	95.0	364	13	US-10-086-623-38	Sequence 38, Appl	619	19	95.0	392	15	US-10-369-493-15069	Sequence 15069, A
547	19	95.0	364	14	US-10-260-533-38	Sequence 38, Appl	620	19	95.0	393	12	US-10-087-192-1923	Sequence 1923, Ap
548	19	95.0	365	14	US-10-156-761-8947	Sequence 8947, Ap	621	19	95.0	394	9	US-09-934-903-4	Sequence 4, Appl
549	19	95.0	367	12	US-10-425-114-64842	Sequence 64842, A	622	19	95.0	394	9	US-09-934-903-4	Sequence 4, Appl
550	19	95.0	368	12	US-10-424-599-237208	Sequence 237208, A	623	19	95.0	394	10	US-09-941-947A-8	Sequence 8, Appl
551	19	95.0	368	15	US-10-369-493-145	Sequence 145, App	624	19	95.0	394	12	US-10-282-122A-54557	Sequence 54557, A
552	19	95.0	368	15	US-10-369-493-2742	Sequence 2742, Ap	625	19	95.0	394	12	US-10-425-114-50770	Sequence 50770, A
553	19	95.0	369	12	US-10-282-122A-56898	Sequence 56898, A	626	19	95.0	394	12	US-10-700-003-4	Sequence 4

673	19	95.0	413	12	US-10-282-122A-48732	Sequence 48732, A	746	19	95.0	439	12	US-10-282-122A-55457	Sequence 55457, A
674	19	95.0	413	15	US-10-369-493-1140	Sequence 1140, Ap	747	19	95.0	439	12	US-10-282-122A-77845	Sequence 77845, A
675	19	95.0	414	9	US-09-820-993-69	Sequence 69, Appl	748	19	95.0	440	12	US-10-282-122A-57263	Sequence 57263, A
676	19	95.0	414	12	US-10-424-599-285027	Sequence 285027, A	749	19	95.0	441	12	US-10-424-599-161614	Sequence 161614, A
677	19	95.0	414	12	US-10-607-565-69	Sequence 69, Appl	750	19	95.0	441	12	US-10-424-599-161615	Sequence 161615, A
678	19	95.0	414	16	US-10-287-226-440	Sequence 440, Appl	751	19	95.0	441	12	US-10-424-599-226388	Sequence 226388, A
679	19	95.0	415	12	US-10-282-122A-65349	Sequence 65349, A	752	19	95.0	444	12	US-10-282-122A-56949	Sequence 56949, A
680	19	95.0	415	12	US-10-282-122A-65787	Sequence 65787, A	753	19	95.0	444	12	US-10-369-493-124	Sequence 124, Appl
681	19	95.0	415	12	US-10-282-122A-68742	Sequence 68742, A	754	19	95.0	445	11	US-09-981-151A-73	Sequence 73, Appl
682	19	95.0	416	14	US-10-156-761-12502	Sequence 1502, A	755	19	95.0	445	12	US-10-282-122A-45092	Sequence 45092, A
683	19	95.0	416	15	US-10-369-493-4737	Sequence 4737, Ap	756	19	95.0	446	9	US-09-798-029-17	Sequence 17, Appl
684	19	95.0	416	15	US-10-369-493-4737	Sequence 4737, Ap	757	19	95.0	446	9	US-10-282-122A-48392	Sequence 48392, A
685	19	95.0	416	15	US-10-369-493-4737	Sequence 4737, Ap	758	19	95.0	446	12	US-10-282-122A-48392	Sequence 52177, A
686	19	95.0	417	12	US-10-424-599-274599	Sequence 274599, A	759	19	95.0	446	12	US-10-425-114-52177	Sequence 52177, A
687	19	95.0	417	12	US-10-425-114-56059	Sequence 56059, A	760	19	95.0	447	9	US-09-908-928-4	Sequence 4, Appl
688	19	95.0	417	14	US-10-128-714-8345	Sequence 345, Ap	761	19	95.0	447	9	US-09-972-528-4	Sequence 72069, A
689	19	95.0	417	14	US-10-128-714-8345	Sequence 845, Ap	762	19	95.0	447	12	US-10-282-122A-72069	Sequence 72069, A
690	19	95.0	418	11	US-09-981-151A-29	Sequence 29, Appl	763	19	95.0	447	12	US-10-282-122A-72107	Sequence 72107, A
691	19	95.0	419	15	US-10-369-493-20475	Sequence 20475, A	764	19	95.0	447	12	US-10-164-966-17	Sequence 17, Appl
692	19	95.0	420	15	US-10-369-493-13961	Sequence 13961, A	765	19	95.0	447	14	US-10-156-761-12559	Sequence 12559, A
693	19	95.0	420	9	US-09-815-442-5303	Sequence 5303, Ap	766	19	95.0	448	12	US-10-282-122A-48951	Sequence 48951, A
694	19	95.0	421	9	US-09-815-442-12388	Sequence 12388, A	767	19	95.0	448	12	US-10-282-122A-48951	Sequence 266964, A
695	19	95.0	421	9	US-09-815-442-12388	Sequence 12388, A	768	19	95.0	448	14	US-10-238-075-1420	Sequence 1420, Ap
696	19	95.0	421	12	US-10-282-122A-71223	Sequence 71223, A	769	19	95.0	449	12	US-10-425-114-39657	Sequence 39657, A
697	19	95.0	421	12	US-10-425-114-56431	Sequence 56431, A	770	19	95.0	450	15	US-10-369-493-2490	Sequence 2490, Ap
698	19	95.0	421	14	US-10-138-701-8	Sequence 8, Appl	771	19	95.0	453	9	US-09-738-626-5635	Sequence 5635, Ap
699	19	95.0	421	14	US-10-138-701-22	Sequence 22, Appl	772	19	95.0	453	12	US-10-282-122A-72044	Sequence 72044, A
700	19	95.0	422	12	US-10-282-122A-66702	Sequence 66702, A	773	19	95.0	454	12	US-10-282-122A-74651	Sequence 74651, A
701	19	95.0	423	12	US-10-282-122A-53283	Sequence 53283, A	774	19	95.0	454	12	US-10-282-122A-74651	Sequence 63654, A
702	19	95.0	423	12	US-10-424-599-204268	Sequence 204268, A	775	19	95.0	454	12	US-10-425-114-63654	Sequence 72, Appl
703	19	95.0	423	12	US-10-425-114-46740	Sequence 46740, A	776	19	95.0	456	12	US-10-282-122A-55920	Sequence 55920, A
704	19	95.0	423	12	US-10-107-431-49	Sequence 49, Appl	777	19	95.0	456	12	US-10-282-122A-78329	Sequence 78329, A
705	19	95.0	423	12	US-10-183-687-222	Sequence 222, Ap	778	19	95.0	456	15	US-10-320-797-3204	Sequence 3204, Ap
706	19	95.0	423	15	US-10-369-493-5557	Sequence 5557, A	779	19	95.0	457	9	US-09-888-615-110	Sequence 110, Appl
707	19	95.0	424	12	US-10-282-122A-45014	Sequence 45014, A	780	19	95.0	457	11	US-09-981-151A-71	Sequence 71, Appl
708	19	95.0	424	12	US-10-424-599-214840	Sequence 214840, A	781	19	95.0	457	12	US-10-282-122A-67190	Sequence 67190, A
709	19	95.0	424	14	US-10-156-761-7776	Sequence 7776, Ap	782	19	95.0	457	14	US-10-156-761-11073	Sequence 11073, A
710	19	95.0	425	14	US-10-156-761-7776	Sequence 7776, Ap	783	19	95.0	457	14	US-10-156-761-13584	Sequence 13584, A
711	19	95.0	425	14	US-10-214-446-20	Sequence 20, Appl	784	19	95.0	458	12	US-10-369-493-3104	Sequence 3104, Ap
712	19	95.0	426	12	US-10-425-114-48506	Sequence 48506, A	785	19	95.0	458	12	US-10-282-122A-45167	Sequence 45167, A
713	19	95.0	427	15	US-10-369-493-2977	Sequence 2977, Ap	786	19	95.0	459	12	US-10-425-114-56969	Sequence 56969, A
714	19	95.0	427	15	US-10-369-493-11890	Sequence 11890, A	787	19	95.0	460	15	US-10-369-493-8005	Sequence 8005, Ap
715	19	95.0	427	15	US-10-104-047-2362	Sequence 2362, Ap	788	19	95.0	460	15	US-10-108-260A-2540	Sequence 2540, Ap
716	19	95.0	428	12	US-10-403-571-110	Sequence 110, Appl	789	19	95.0	461	12	US-10-282-122A-50440	Sequence 50440, Ap
717	19	95.0	428	15	US-10-369-493-17335	Sequence 17335, A	790	19	95.0	464	15	US-10-259-194A-186	Sequence 186, Appl
718	19	95.0	431	12	US-10-282-122A-47233	Sequence 47233, A	791	19	95.0	465	12	US-10-424-599-266524	Sequence 266524, A
719	19	95.0	431	12	US-10-282-122A-51920	Sequence 51920, A	792	19	95.0	465	14	US-10-193-144-19	Sequence 19, Appl
720	19	95.0	431	12	US-10-424-599-212351	Sequence 212351, A	793	19	95.0	465	15	US-10-345-072-19	Sequence 19, Appl
721	19	95.0	431	12	US-10-041-018-260	Sequence 260, Appl	794	19	95.0	466	12	US-10-282-122A-63275	Sequence 63275, A
722	19	95.0	431	15	US-10-369-493-21530	Sequence 21530, A	795	19	95.0	466	15	US-10-120-801-61	Sequence 61, Appl
723	19	95.0	432	12	US-10-425-114-72646	Sequence 72646, A	796	19	95.0	468	15	US-10-369-493-1416	Sequence 1416, Ap
724	19	95.0	432	15	US-10-369-493-8573	Sequence 8573, Ap	797	19	95.0	469	12	US-10-417-700A-65	Sequence 65, Appl
725	19	95.0	433	14	US-10-128-714-8545	Sequence 8545, Ap	798	19	95.0	469	12	US-10-417-700A-77	Sequence 77, Appl
726	19	95.0	433	15	US-10-369-493-10203	Sequence 10203, A	799	19	95.0	469	14	US-10-034-585-7098	Sequence 7098, Ap
727	19	95.0	434	12	US-10-282-122A-66904	Sequence 66904, A	800	19	95.0	471	12	US-10-282-122A-76782	Sequence 76782, A
728	19	95.0	435	12	US-10-627-476-238	Sequence 238, Appl	801	19	95.0	471	12	US-10-425-114-72794	Sequence 72794, A
729	19	95.0	436	12	US-10-282-122A-50291	Sequence 50291, A	802	19	95.0	472	11	US-09-981-151A-27	Sequence 27, Appl
730	19	95.0	436	12	US-10-425-114-46821	Sequence 46821, A	803	19	95.0	472	12	US-10-424-599-251597	Sequence 251597, A
731	19	95.0	436	12	US-10-041-018-235	Sequence 235, Appl	804	19	95.0	472	12	US-10-425-114-62714	Sequence 62714, A
732	19	95.0	436	15	US-10-369-493-21378	Sequence 21378, A	805	19	95.0	474	12	US-10-282-122A-82083	Sequence 82083, A
733	19	95.0	437	9	US-09-919-932-2	Sequence 2, Appl	806	19	95.0	474	12	US-10-425-114-40744	Sequence 40744, A
734	19	95.0	437	9	US-09-738-626-4426	Sequence 426, Ap	807	19	95.0	475	12	US-10-425-114-39831	Sequence 39831, A
735	19	95.0	437	12	US-09-746-660A-2	Sequence 2, Appl	808	19	95.0	475	12	US-10-424-599-257731	Sequence 257731, A
736	19	95.0	437	12	US-10-282-122A-42940	Sequence 42940, A	809	19	95.0	476	12	US-10-156-761-14307	Sequence 14307, A
737	19	95.0	437	12	US-10-282-122A-73421	Sequence 73421, A	810	19	95.0	478	12	US-10-425-114-38072	Sequence 38072, A
738	19	95.0	437	12	US-10-450-055-6	Sequence 6, Appl	811	19	95.0	478	9	US-09-820-893-108	Sequence 108, Appl
739	19	95.0	437	15	US-10-177-293-4180	Sequence 457, Ap	812	19	95.0	480	12	US-10-424-599-166523	Sequence 166523, A
740	19	95.0	437	15	US-10-395-027-1180	Sequence 1180, Appl	813	19	95.0	480	12	US-10-425-114-63937	Sequence 63937, A
741	19	95.0	438	12	US-10-282-122A-56153	Sequence 56153, A	814	19	95.0	480	12	US-10-607-565-108	Sequence 108, Appl
742	19	95.0	438	12	US-10-282-122A-56161	Sequence 56161, A	815	19	95.0	481	12	US-10-424-599-166520	Sequence 166520, A
743	19	95.0	438	12	US-10-282-122A-75202	Sequence 75202, A	816	19	95.0	481	12	US-10-424-599-166520	Sequence 48208, A
744	19	95.0	438	12	US-10-282-122A-75609	Sequence 75609, A	817	19	95.0	483	12	US-10-282-122A-48208	Sequence 48208, A
745	19	95.0	439	11	US-09-764-875-930	Sequence 930, Appl	818	19	95.0	484	12	US-10-425-114-42223	Sequence 42223, A

819	19	95.0	485	9	US-09-769-864-1	Sequence 1, Appli	892	19	95.0	518	12	US-10-424-599-172544	Sequence 172544,
820	19	95.0	485	9	US-09-769-864-2	Sequence 2, Appli	893	19	95.0	522	15	US-10-099-322-134	Sequence 124, App
821	19	95.0	485	9	US-09-769-864-6	Sequence 6, Appli	894	19	95.0	522	15	US-10-099-322-125	Sequence 125, App
822	19	95.0	485	9	US-09-769-864-7	Sequence 7, Appli	895	19	95.0	522	15	US-10-044-564-124	Sequence 124, App
823	19	95.0	485	9	US-09-769-864-8	Sequence 8, Appli	896	19	95.0	522	15	US-10-044-564-125	Sequence 125, App
824	19	95.0	485	9	US-09-854-346-2	Sequence 2, Appli	897	19	95.0	523	15	US-10-099-322-122	Sequence 122, App
825	19	95.0	485	9	US-09-854-346-4	Sequence 4, Appli	898	19	95.0	523	15	US-10-099-322-123	Sequence 123, App
826	19	95.0	485	9	US-09-854-346-12	Sequence 12, Appli	899	19	95.0	523	15	US-10-044-564-122	Sequence 122, App
827	19	95.0	485	9	US-09-854-346-13	Sequence 13, Appli	900	19	95.0	523	15	US-10-044-564-123	Sequence 123, App
828	19	95.0	485	9	US-09-902-188A-1	Sequence 1, Appli	901	19	95.0	524	12	US-10-425-114-62567	Sequence 62567, A
829	19	95.0	485	9	US-09-902-188A-1	Sequence 2, Appli	902	19	95.0	524	12	US-10-425-114-62567	Sequence 62567, A
830	19	95.0	485	9	US-09-902-188A-7	Sequence 7, Appli	903	19	95.0	524	12	US-10-425-114-62567	Sequence 62567, A
831	19	95.0	485	9	US-09-918-543-2	Sequence 2, Appli	904	19	95.0	524	12	US-10-403-161-34	Sequence 34, Appli
832	19	95.0	485	9	US-09-918-543-4	Sequence 4, Appli	905	19	95.0	524	12	US-10-403-161-40	Sequence 40, Appli
833	19	95.0	485	9	US-09-918-543-12	Sequence 12, Appli	906	19	95.0	524	12	US-10-403-161-42	Sequence 42, Appli
834	19	95.0	485	9	US-09-918-543-12	Sequence 12, Appli	907	19	95.0	524	15	US-10-099-322-34	Sequence 34, Appli
835	19	95.0	485	9	US-09-795-211-1	Sequence 1, Appli	908	19	95.0	524	15	US-10-099-322-121	Sequence 121, App
836	19	95.0	485	9	US-09-795-211-2	Sequence 2, Appli	909	19	95.0	524	15	US-10-044-564-34	Sequence 34, Appli
837	19	95.0	485	10	US-09-925-576C-2	Sequence 2, Appli	910	19	95.0	524	9	US-09-815-242-13696	Sequence 13696, A
838	19	95.0	485	10	US-09-925-576C-4	Sequence 4, Appli	911	19	95.0	526	12	US-10-282-122A-73677	Sequence 73677, A
839	19	95.0	485	10	US-09-925-576C-13	Sequence 13, Appli	912	19	95.0	526	12	US-10-282-122A-73677	Sequence 73677, A
840	19	95.0	485	12	US-10-665-667-1	Sequence 1, Appli	913	19	95.0	526	12	US-10-425-114-45814	Sequence 45814, A
841	19	95.0	485	12	US-10-665-667-2	Sequence 2, Appli	914	19	95.0	527	12	US-10-425-114-45814	Sequence 45814, A
842	19	95.0	485	12	US-10-665-667-6	Sequence 6, Appli	915	19	95.0	529	12	US-10-425-114-45814	Sequence 45814, A
843	19	95.0	485	12	US-10-665-667-7	Sequence 7, Appli	916	19	95.0	529	12	US-10-403-161-38	Sequence 38, Appli
844	19	95.0	485	12	US-10-665-667-8	Sequence 8, Appli	917	19	95.0	530	15	US-10-369-493-19139	Sequence 19139, A
845	19	95.0	485	12	US-10-025-648-1	Sequence 1, Appli	918	19	95.0	531	12	US-10-369-493-19139	Sequence 19139, A
846	19	95.0	485	12	US-10-025-648-2	Sequence 2, Appli	919	19	95.0	531	12	US-10-425-114-45814	Sequence 45814, A
847	19	95.0	485	12	US-10-025-648-7	Sequence 7, Appli	920	19	95.0	533	12	US-10-282-122A-61253	Sequence 61253, A
848	19	95.0	485	12	US-10-025-648-7	Sequence 7, Appli	921	19	95.0	533	14	US-10-156-761-9551	Sequence 9551, Ap
849	19	95.0	485	12	US-10-327-837-1	Sequence 1, Appli	922	19	95.0	535	12	US-10-282-122A-60086	Sequence 60086, A
850	19	95.0	485	12	US-10-327-837-2	Sequence 2, Appli	923	19	95.0	535	12	US-10-425-114-63263	Sequence 63263, A
851	19	95.0	485	12	US-10-327-837-6	Sequence 6, Appli	924	19	95.0	536	12	US-10-282-122A-74974	Sequence 74974, A
852	19	95.0	485	12	US-10-327-837-7	Sequence 7, Appli	925	19	95.0	536	12	US-10-425-114-69007	Sequence 69007, A
853	19	95.0	485	12	US-10-327-837-8	Sequence 8, Appli	926	19	95.0	536	15	US-10-369-493-7664	Sequence 7664, Ap
854	19	95.0	485	12	US-10-327-837-24	Sequence 24, Appli	927	19	95.0	537	12	US-10-282-122A-75618	Sequence 75618, A
855	19	95.0	485	12	US-10-327-837-26	Sequence 26, Appli	928	19	95.0	537	12	US-10-425-114-64000	Sequence 64000, A
856	19	95.0	485	14	US-10-209-812-5	Sequence 5, Appli	929	19	95.0	537	15	US-10-120-801-12	Sequence 12, Appli
857	19	95.0	485	14	US-10-184-771-12	Sequence 12, Appli	930	19	95.0	538	12	US-10-282-122A-44952	Sequence 44952, A
858	19	95.0	485	15	US-10-369-493-4906	Sequence 4906, Ap	931	19	95.0	538	12	US-10-425-114-70508	Sequence 70508, A
859	19	95.0	487	12	US-10-403-161-44	Sequence 44, Appli	932	19	95.0	539	12	US-10-282-122A-57265	Sequence 57265, A
860	19	95.0	487	15	US-10-369-493-18040	Sequence 18040, A	933	19	95.0	539	14	US-10-425-114-46493	Sequence 46493, A
861	19	95.0	490	12	US-10-425-114-45580	Sequence 45580, A	934	19	95.0	539	14	US-10-156-761-11731	Sequence 11731, A
862	19	95.0	491	10	US-09-863-776-34	Sequence 34, Appli	935	19	95.0	539	14	US-10-230-026-34	Sequence 34, Appli
863	19	95.0	491	12	US-09-881-752A-142	Sequence 142, App	936	19	95.0	540	9	US-09-738-626-4450	Sequence 4450, Ap
864	19	95.0	492	12	US-10-282-122A-58846	Sequence 58846, A	937	19	95.0	540	15	US-10-369-493-3921	Sequence 3921, Ap
865	19	95.0	493	9	US-09-918-911-5	Sequence 5, Appli	938	19	95.0	543	12	US-10-282-122A-77151	Sequence 77151, A
866	19	95.0	493	12	US-10-424-599-284218	Sequence 284218, A	939	19	95.0	543	15	US-10-369-493-10792	Sequence 10792, A
867	19	95.0	493	12	US-10-614-076-70	Sequence 70, Appli	940	19	95.0	546	12	US-10-087-193-843	Sequence 843, App
868	19	95.0	494	15	US-10-369-493-20452	Sequence 20452, A	941	19	95.0	547	12	US-10-282-122A-53681	Sequence 53681, A
869	19	95.0	496	12	US-10-425-114-58882	Sequence 58882, A	942	19	95.0	547	14	US-10-043-344-149	Sequence 149, App
870	19	95.0	497	15	US-10-369-493-5803	Sequence 5803, Ap	943	19	95.0	548	15	US-10-282-122A-65960	Sequence 65960, A
871	19	95.0	498	15	US-10-369-493-11919	Sequence 11919, A	944	19	95.0	548	15	US-10-369-493-6496	Sequence 6496, Ap
872	19	95.0	499	9	US-09-815-242-11035	Sequence 11035, A	945	19	95.0	549	9	US-09-815-242-11206	Sequence 11206, A
873	19	95.0	499	12	US-10-282-122A-58172	Sequence 58172, A	946	19	95.0	549	12	US-10-282-122A-58451	Sequence 58451, A
874	19	95.0	499	12	US-10-424-599-149746	Sequence 149746, A	947	19	95.0	549	12	US-10-282-122A-66993	Sequence 66993, A
875	19	95.0	499	15	US-10-369-493-8665	Sequence 8665, Ap	948	19	95.0	549	15	US-10-282-122A-66993	Sequence 66993, A
876	19	95.0	501	12	US-10-282-122A-63359	Sequence 63359, A	949	19	95.0	550	11	US-10-369-493-5478	Sequence 5478, Ap
877	19	95.0	503	12	US-10-425-114-56039	Sequence 56039, A	950	19	95.0	550	12	US-10-369-493-5478	Sequence 5478, Ap
878	19	95.0	504	9	US-09-738-626-7001	Sequence 7001, Ap	951	19	95.0	550	12	US-10-336-049-2	Sequence 2, Appli
879	19	95.0	504	12	US-10-424-599-162427	Sequence 162427, A	952	19	95.0	550	12	US-10-425-114-41988	Sequence 41988, A
880	19	95.0	504	12	US-10-627-476-500	Sequence 500, App	953	19	95.0	550	14	US-10-091-322-2	Sequence 2, Appli
881	19	95.0	504	12	US-10-425-114-64036	Sequence 64036, A	954	19	95.0	552	12	US-10-282-122A-59108	Sequence 59108, A
882	19	95.0	505	15	US-10-369-493-8682	Sequence 8682, Ap	955	19	95.0	552	12	US-10-282-122A-70501	Sequence 70501, A
883	19	95.0	506	9	US-09-943-671-9	Sequence 9, Appli	956	19	95.0	553	12	US-10-282-122A-71716	Sequence 71716, A
884	19	95.0	506	15	US-10-405-877-95	Sequence 95, Appli	957	19	95.0	553	15	US-10-313-371A-2	Sequence 2, Appli
885	19	95.0	506	15	US-10-405-877-95	Sequence 95, Appli	958	19	95.0	553	15	US-10-313-371A-4	Sequence 4, Appli
886	19	95.0	515	9	US-09-925-300-1285	Sequence 1285, Ap	959	19	95.0	556	15	US-10-369-493-5442	Sequence 5442, Ap
887	19	95.0	515	12	US-10-424-599-146129	Sequence 146129, A	960	19	95.0	556	12	US-10-369-493-11974	Sequence 11974, A
888	19	95.0	515	12	US-10-425-114-68917	Sequence 68917, A	961	19	95.0	558	12	US-10-363-616-374	Sequence 374, App
889	19	95.0	516	9	US-09-986-676A-2	Sequence 2, Appli	962	19	95.0	558	15	US-10-369-493-6382	Sequence 6382, Ap
890	19	95.0	516	9	US-09-971-611-2	Sequence 2, Appli	963	19	95.0	559	9	US-09-925-300-1479	Sequence 1479, Ap
891	19	95.0	516	16	US-10-399-161-8	Sequence 8, Appli	964	19	95.0	559	12	US-10-282-122A-65056	Sequence 65056, A

965	19	95.0	559	12	US-10-282-122A-75373	Sequence 73373, A
966	19	95.0	559	12	US-10-282-122A-75373	Sequence 73373, A
967	19	95.0	561	12	US-10-282-122A-75352	Sequence 73552, A
968	19	95.0	561	12	US-10-282-122A-65931	Sequence 65931, A
969	19	95.0	564	16	US-10-369-493-21529	Sequence 21529, A
970	19	95.0	565	12	US-10-389-566-1969	Sequence 1969, Ap
971	19	95.0	565	12	US-10-380-565-34	Sequence 34, Appl
972	19	95.0	566	14	US-10-100-957A-142	Sequence 142, App
973	19	95.0	566	14	US-10-111-088-19	Sequence 19, Appl
974	19	95.0	567	12	US-10-425-114-45750	Sequence 45750, A
975	19	95.0	568	12	US-10-282-122A-65302	Sequence 65302, A
976	19	95.0	573	9	US-09-764-868-671	Sequence 671, App
977	19	95.0	573	11	US-09-764-875-1205	Sequence 1205, Ap
978	19	95.0	574	12	US-10-425-114-40721	Sequence 40721, A
979	19	95.0	575	12	US-10-425-114-42248	Sequence 42248, A
980	19	95.0	575	12	US-10-424-599-18618	Sequence 18618, A
981	19	95.0	575	15	US-10-312-352-8	Sequence 8, Appl
982	19	95.0	576	15	US-10-120-801-60	Sequence 60, Appl
983	19	95.0	577	9	US-09-815-942-11305	Sequence 11305, Ap
984	19	95.0	577	12	US-10-282-122A-58657	Sequence 58657, A
985	19	95.0	577	12	US-10-335-977-7019	Sequence 7019, Ap
986	19	95.0	580	9	US-09-841-132-598	Sequence 598, App
987	19	95.0	580	12	US-10-282-122A-54815	Sequence 54815, A
988	19	95.0	580	15	US-10-389-762-332	Sequence 332, App
989	19	95.0	581	12	US-10-424-599-255268	Sequence 255268, A
990	19	95.0	584	15	US-10-108-260A-4380	Sequence 4380, Ap
991	19	95.0	585	12	US-10-431-449-2	Sequence 2, Appl
992	19	95.0	585	12	US-10-431-449-6	Sequence 6, Appl
993	19	95.0	585	12	US-10-431-449-6	Sequence 6, Appl
994	19	95.0	587	12	US-10-424-599-198064	Sequence 198064, A
995	19	95.0	587	13	US-10-042-523-3	Sequence 3, Appl
996	19	95.0	588	15	US-10-309-560-10	Sequence 10, Appl
997	19	95.0	590	12	US-10-282-122A-65911	Sequence 65911, A
998	19	95.0	592	12	US-10-424-599-214104	Sequence 214104, A
999	19	95.0	592	12	US-10-335-977-5869	Sequence 5869, Ap
1000	19	95.0	595	12	US-10-382-122A-44744	Sequence 44744, A

## ALIGNMENTS

```

1  RESULT 1
2  JS-09-791-378-343
3  ; Sequence 343, Application US/09791378
4  ; Patent No. US20020142303A1
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Parekh, Rajesh
7  ; TITLE OF INVENTION: PROTEINS, GENES AND THEIR USE FOR DIAGNOSIS AND TREATMENT OF
8  ; TITLE OF INVENTION: SCHIZOPHRENIA
9  ; FILE REFERENCE: 9195-061-999
10 ; CURRENT APPLICATION NUMBER: US/09/791,378
11 ; CURRENT FILING DATE: 2001-02-23
12 ; PRIOR APPLICATION NUMBER: 09/750,395
13 ; PRIOR FILING DATE: 2000-12-28
14 ; NUMBER OF SEQ ID NOS: 677
15 ; SOFTWARE: Patent in version 3.0
16 ; SEQ ID NO 343
17 ; LENGTH: 8
18 ; TYPE: PRT
19 ; ORGANISM: Homo sapiens
20 ; JS-09-791-378-343

```

```

; Sequence 206, Application US/09791393
;
; GENERAL INFORMATION:
;
; APPLICANT: Herath, Herath Mudiyansele Athula Chandrasiri
; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Rohloff, Christian
;
; TITLE OF INVENTION: Proteins, Genes and Their Use for
;
; TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
;
; TITLE OF INVENTION: and Unipolar Depression
;
; FILE REFERENCE: 2543-1-001 N1
;
; CURRENT APPLICATION NUMBER: US/09791393
;
; CURRENT FILING DATE: 2002-01-02
;
; EARLIER APPLICATION NUMBER: GB 0004412.3
;
; EARLIER FILING DATE: 2000-02-24
;
; EARLIER APPLICATION NUMBER: GB 0030050.9
;
; EARLIER FILING DATE: 2000-12-08
;
; EARLIER APPLICATION NUMBER: US 60/254,830
;
; EARLIER FILING DATE: 2000-12-12
;
; NUMBER OF SEQ ID NOS: 308
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 206
;
; LENGTH: 8
;
; TYPE: PRT
;
; ORGANISM: homo sapien
;
US-09-791-393-206

Query Match          95.0%; Score 19; DB 10; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Query Match      95.0%; Score 19; DB 10; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 VAEF 5
        ||||
Db      2 VAEF 5

```



SULT 4  
 -10-128-711-86  
 Sequence 86, Application US/10128711  
 Publication No. US20030099634A1  
 GENERAL INFORMATION:  
 APPLICANT: VITIELLO, Maria A.  
 CHESTNUT, Robert W.  
 SETTE, Alessandro D.  
 CELIS, Esteban  
 GRAY, Howard

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
 CYT IMMUNITY

NUMBER OF SEQUENCES: 153  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend Kourie and Crew  
 STREET: Steuart Street Tower, One Market Plaza  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: US  
 ZIP: 94105-1493

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/128,711  
 FILING DATE: 22-Apr-2002  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/197,484  
 FILING DATE: 16-FEB-1994  
 APPLICATION NUMBER: US 07/935,811  
 FILING DATE: 26-AUG-1992  
 APPLICATION NUMBER: US 07/874,491  
 FILING DATE: 27-APR-1992  
 APPLICATION NUMBER: US 07/827,682  
 FILING DATE: 29-JAN-1992  
 APPLICATION NUMBER: US 07/749,568  
 FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:  
 NAME: Parmelee, Steven W.  
 REGISTRATION NUMBER: 31,990  
 REFERENCE/DOCKET NUMBER: 14137-26-4

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 467-9600  
 TELEFAX: (206) 623-6793

INFORMATION FOR SEQ ID NO: 86:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 9 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown

MOLECULE TYPE: peptide  
 SEQUENCE DESCRIPTION: SEQ ID NO: 86:  
 -10-128-711-86

Query Match 95.0%; Score 19; DB 14; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
 ||||  
 2 VAEF 5

SULT 5  
 -10-128-711-145  
 Sequence 145, Application US/10128711  
 Publication No. US20030099634A1  
 GENERAL INFORMATION:  
 APPLICANT: VITIELLO, Maria A.

CHESTNUT, Robert W.  
 SETTE, Alessandro D.  
 CELIS, Esteban  
 GRAY, Howard

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
 CYT IMMUNITY

NUMBER OF SEQUENCES: 153  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend Kourie and Crew  
 STREET: Steuart Street Tower, One Market Plaza  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: US  
 ZIP: 94105-1493

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/128,711  
 FILING DATE: 22-Apr-2002  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/197,484  
 FILING DATE: 16-FEB-1994  
 APPLICATION NUMBER: US 07/935,811  
 FILING DATE: 26-AUG-1992  
 APPLICATION NUMBER: US 07/874,491  
 FILING DATE: 27-APR-1992  
 APPLICATION NUMBER: US 07/827,682  
 FILING DATE: 29-JAN-1992  
 APPLICATION NUMBER: US 07/749,568  
 FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:  
 NAME: Parmelee, Steven W.  
 REGISTRATION NUMBER: 31,990  
 REFERENCE/DOCKET NUMBER: 14137-26-4

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 467-9600  
 TELEFAX: (206) 623-6793

INFORMATION FOR SEQ ID NO: 145:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 9 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown

MOLECULE TYPE: peptide  
 SEQUENCE DESCRIPTION: SEQ ID NO: 145:  
 US-10-128-711-145

Query Match 95.0%; Score 19; DB 14; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
 ||||  
 2 VAEF 5

RESULT 6  
 US-09-508-943A-197  
 ; Sequence 197, Application US/09908943A  
 ; Publication No. US20030017991A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yan, Riqiang  
 ; APPLICANT: Tomasselli, Alfredo G.  
 ; APPLICANT: Gurney, Mark E.  
 ; APPLICANT: Emmons, Thomas L.  
 ; APPLICANT: Bienkowski, Mike J.  
 ; APPLICANT: Heinrikson, Robert L.  
 ; TITLE OF INVENTION: SUBSTRATES AND ASSAYS FOR BETA-SECRETASE ACTIVITY  
 ; FILE REFERENCE: 29915/00281A.US1



CURRENT APPLICATION NUMBER: US/09/908,943A  
CURRENT FILING DATE: 2001-07-19  
PRIOR APPLICATION NUMBER: 60/219,795  
PRIOR FILING DATE: 2000-07-19  
NUMBER OF SEQ ID NOS: 197  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 197  
LENGTH: 10  
TYPE: PRT  
ORGANISM: synthetic peptide sequence  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (4)..(4)  
OTHER INFORMATION: amino acid at position 4 has been derivatized with a statine  
NAME/KEY: SITE  
LOCATION: (10)..(10)  
OTHER INFORMATION: amino acid at position 10 has been derivatized with Bodipy FL  
JS-09-908-943A-197

Query Match 95.0%; Score 19; DB 12; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|  
|  
|  
|  
DB 5 VAEF 8

RESULT 7  
JS-09-908-943A-196  
Sequence 196, Application US/0908943A  
Publication No. US20030017991A1  
GENERAL INFORMATION:  
APPLICANT: Yan, Riqiang  
APPLICANT: Tonasselli, Alfredo G.  
APPLICANT: Gurney, Mark E.  
APPLICANT: Emmons, Thomas L.  
APPLICANT: Bienkowski, Mike J.  
APPLICANT: Heinrikson, Robert L.  
TITLE OF INVENTION: SUBSTRATES AND ASSAYS FOR BETA-SECRETASE ACTIVITY  
FILE REFERENCE: 29915/00281A.US1  
CURRENT APPLICATION NUMBER: US/09/908,943A  
CURRENT FILING DATE: 2001-07-19  
PRIOR APPLICATION NUMBER: 60/219,795  
PRIOR FILING DATE: 2000-07-19  
NUMBER OF SEQ ID NOS: 197  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 196  
LENGTH: 12  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Description of artificial sequence: synthetic peptide sequence  
NAME/KEY: SITE  
LOCATION: (4)..(4)  
OTHER INFORMATION: amino acid at position 4 has been derivatized with a statine  
JS-09-908-943A-196

Query Match 95.0%; Score 19; DB 12; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|  
|  
|  
|  
DB 5 VAEF 8

RESULT 8  
JS-10-032-818-34  
Sequence 34, Application US/10032818  
Publication No. US20030092629A1  
GENERAL INFORMATION:  
APPLICANT: Tang, Jordan J.N.

APPLICANT: Koelsch, Gerald  
APPLICANT: Ghosh, Arun K.  
TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof  
FILE REFERENCE: 2932.1006-007  
CURRENT APPLICATION NUMBER: US/10/032,818  
CURRENT FILING DATE: 2001-12-28  
PRIOR APPLICATION NUMBER: US 60/275,756  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: US 60/258,705  
PRIOR FILING DATE: 2000-12-28  
NUMBER OF SEQ ID NOS: 83  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 34  
LENGTH: 12  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Synthetic peptide  
US-10-032-818-34

Query Match 95.0%; Score 19; DB 14; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|  
|  
|  
|  
DB 5 VAEF 8

RESULT 9  
US-10-239-313A-379  
Sequence 379, Application US/10239313A  
Publication No. US20030175285A1  
GENERAL INFORMATION:  
APPLICANT: KLINGUER - HAMOUR, Christine  
APPLICANT: CORVAIA, Nathalie  
APPLICANT: BECK, Alain  
APPLICANT: GORTSCH, Liliane  
TITLE OF INVENTION: MOLECULE OF PHARMACEUTICAL INTEREST COMPRISING AT ITS N-TERMINAL A GLUTAMIC ACID OR A GLUTAMINE IN THE FORM OF A PHYSIOLOGICALLY ACCEPTABLE STRONG ACID  
FILE REFERENCE: 343 727 - US  
CURRENT APPLICATION NUMBER: US/10/239,313A  
CURRENT FILING DATE: 2002-09-19  
PRIOR APPLICATION NUMBER: FR 00/03711  
PRIOR FILING DATE: 2000-03-23  
PRIOR APPLICATION NUMBER: PCT 01/70772  
PRIOR FILING DATE: 2001-03-22  
NUMBER OF SEQ ID NOS: 697  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 379  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Mycobacterium leprae  
US-10-239-313A-379

Query Match 95.0%; Score 19; DB 14; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
|  
|  
|  
|  
DB 4 VAEF 7

RESULT 10  
US-10-119-528-7  
Sequence 7, Application US/10119528  
Publication No. US20030175722A1  
GENERAL INFORMATION:  
APPLICANT: Mann, M.  
APPLICANT: Mortensen, P.  
TITLE OF INVENTION: METHODS AND SYSTEMS FOR SEARCHING GENOMIC DATABASES

FILE REFERENCE: MDSP-P01-004  
CURRENT APPLICATION NUMBER: US/10/119,528  
CURRENT FILING DATE: 2002-04-09  
PRIOR APPLICATION NUMBER: 60/282,551  
PRIOR FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: 60/285,362  
PRIOR FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 7  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
S-10-119-528-7

Query Match 95.0%; Score 19; DB 14; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
|||  
b 10 VAEF 13

RESULT 11  
S-09-171-432A-62  
Sequence 62, Application US/09171432A  
Publication No. US20030187184A1  
GENERAL INFORMATION:  
APPLICANT: Fields, Howard A.  
APPLICANT: Khudyakov, Yuri E.  
TITLE OF INVENTION: Antigenically Reactive Regions of the  
TITLE OF INVENTION: Hepatitis A Virus Polyprotein  
NUMBER OF SEQUENCES: 88  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick Stockton LLP  
STREET: 3424 Peachtree Road, N.E.  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30326  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 23-NOV-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Greene, Jamie L.  
REGISTRATION NUMBER: 32,467  
REFERENCE/DOCKET NUMBER: 03063-0231US  
TELEPHONE: (404) 949-2400  
TELEFAX: (404) 949-2499  
INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /label= YK-1368  
US-09-171-432A-62

RESULT 13  
US-09-171-432A-65  
Sequence 65, Application US/09171432A  
Publication No. US20030187184A1  
GENERAL INFORMATION:  
APPLICANT: Fields, Howard A.

Query Match 95.0%; Score 19; DB 10; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAEF 5  
|||  
Db 13 VAEF 16

RESULT 12  
US-09-171-432A-63  
Sequence 63, Application US/09171432A  
Publication No. US20030187184A1  
GENERAL INFORMATION:  
APPLICANT: Fields, Howard A.  
APPLICANT: Khudyakov, Yuri E.  
TITLE OF INVENTION: Antigenically Reactive Regions of the  
TITLE OF INVENTION: Hepatitis A Virus Polyprotein  
NUMBER OF SEQUENCES: 88  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick Stockton LLP  
STREET: 3424 Peachtree Road, N.E.  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30326  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 23-NOV-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Greene, Jamie L.  
REGISTRATION NUMBER: 32,467  
REFERENCE/DOCKET NUMBER: 03063-0231US  
TELEPHONE: (404) 949-2400  
TELEFAX: (404) 949-2499  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /label= YK-1369  
US-09-171-432A-63

Query Match 95.0%; Score 19; DB 10; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAEF 5  
|||  
Db 4 VAEF 7

RESULT 13  
US-09-171-432A-65  
Sequence 65, Application US/09171432A  
Publication No. US20030187184A1  
GENERAL INFORMATION:  
APPLICANT: Fields, Howard A.

APPLICANT: Khuyakov, Yury E.  
TITLE OF INVENTION: Antigenically Reactive Regions of the  
Hepatitis A Virus Polypeptide  
NUMBER OF SEQUENCES: 88  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick Stockton LLP  
STREET: 3424 Peachtree Road, N.E.  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30326  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/171.432A  
FILING DATE: 23-NOV-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Greene, Jamie L.  
REGISTRATION NUMBER: 32,467  
REFERENCE/DOCKET NUMBER: 03063-0231US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 949-2400  
TELEFAX: (404) 949-2499  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..25  
OTHER INFORMATION: /label= YK-1832  
US-09-171-432A-65

Query Match 95.0%; Score 19; DB 10; Length 25;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 3 VAEF 6

RESULT 14  
US-09-864-761-47201  
Sequence 47201, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aemica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
SEQ ID NO 47201  
LENGTH: 28  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC005772.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.46  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.67  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.61  
US-09-864-761-47201

Query Match 95.0%; Score 19; DB 9; Length 28;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 4 VAEF 7

RESULT 15  
US-10-424-599-216297  
Sequence 216297, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J  
APPLICANT: Kovalic, David K  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285694  
SEQ ID NO 216297  
LENGTH: 28  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_37345C.1.pep  
US-10-424-599-216297

```
Query Match          95.0%; Score 19; DB 12; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      2 VAEF 5
      ||||
      10 VAEF 13

RESULT 16
;10-043-344-69
; Sequence 69, Application US/10043344
; Publication No. US2003008806A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Harkness, Robin E.
; APPLICANT: Schryvers, Anthony B.
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Murrin, Andrew D.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: TRANSFERIN RECEPTOR GENES
; FILE REFERENCE: 1038-1221 MIS
; CURRENT APPLICATION NUMBER: US/10/043,344
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 08/649,518
; PRIOR FILING DATE: 1996-05-17
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 69
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
;10-043-344-69

Query Match          95.0%; Score 19; DB 14; Length 35;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      2 VAEF 5
      ||||
      14 VAEF 17

RESULT 17
;10-242-355-570
; Sequence 570, Application US/10242355
; Publication No. US2003023583A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC003C1
; CURRENT APPLICATION NUMBER: US/10/242,355
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,897
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,495
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
```

```
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1267
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 570
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
;10-242-355-570

Query Match          95.0%; Score 19; DB 15; Length 37;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      2 VAEF 5
      ||||
      14 VAEF 17

RESULT 18
;10-430-752A-49
; Sequence 49, Application US/10430752A
; Publication No. US20040005669A1
; GENERAL INFORMATION:
; APPLICANT: Stahl, Stefan
; APPLICANT: Jonasson, Per
; APPLICANT: Nygren, Per-Ake
; APPLICANT: Uhlen, Mathias
; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF INSULIN
; FILE REFERENCE: 11541-003001
; CURRENT APPLICATION NUMBER: US/10/430,752A
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 09/485,286
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: PCT/GB98/02382
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: GB 9716790.2
; PRIOR FILING DATE: 1997-08-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Verasper moseri
;10-430-752A-49

Query Match          95.0%; Score 19; DB 15; Length 37;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      2 VAEF 5
      ||||
      24 VAEF 27

RESULT 19
;10-430-752A-45
; Sequence 45, Application US/10430752A
; Publication No. US20040005669A1
; GENERAL INFORMATION:
; APPLICANT: Stahl, Stefan
; APPLICANT: Jonasson, Per
; APPLICANT: Nygren, Per-Ake
; APPLICANT: Uhlen, Mathias
; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF INSULIN
; FILE REFERENCE: 11541-003001
; CURRENT APPLICATION NUMBER: US/10/430,752A
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 09/485,286
; PRIOR FILING DATE: 2000-02-07
```

PRIOR APPLICATION NUMBER: PCT/GB98/02382  
 PRIOR FILING DATE: 1998-08-07  
 PRIOR APPLICATION NUMBER: GB 9716790.2  
 PRIOR FILING DATE: 1997-08-07  
 NUMBER OF SEQ ID NOS: 63  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 45  
 LENGTH: 38  
 TYPE: PRT  
 ORGANISM: Lophius piscatorius  
 S-10-430-752A-45

Query Match 95.0%; Score 19; DB 15; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 2 VAEF 5  
 ||||  
 b 25 VAEF 28

## RESULT 20

S-10-372-003A-71  
 Sequence 71, Application US/10372003A  
 Publication No. US20030215846A1

## GENERAL INFORMATION:

APPLICANT: Watt, Paul  
 APPLICANT: Thomas, Wayne  
 APPLICANT: Hopkins, Richard  
 TITLE OF INVENTION: Methods of constructing and screening  
 TITLE OF INVENTION: diverse expression libraries  
 FILE REFERENCE: FBRI40.001CPI  
 CURRENT APPLICATION NUMBER: US/10/372,003A  
 CURRENT FILING DATE: 2003-02-21  
 PRIOR APPLICATION NUMBER: US 09/568,229  
 PRIOR FILING DATE: 2000-05-05  
 PRIOR APPLICATION NUMBER: US 60/132,711  
 PRIOR FILING DATE: 1999-05-05  
 NUMBER OF SEQ ID NOS: 81  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 71  
 LENGTH: 41  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: synthetic clone BGF13 with first open reading  
 OTHER INFORMATION: frame

## S-10-372-003A-71

Query Match 95.0%; Score 19; DB 15; Length 41;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 2 VAEF 5  
 ||||  
 b 28 VAEF 31

## RESULT 21

S-09-864-761-37202  
 Sequence 37202, Application US/09864761  
 Patent No. US20020048763A1

## GENERAL INFORMATION:

APPLICANT: Fenn, Sharon G.  
 APPLICANT: Rank, David R.  
 APPLICANT: Hanzel, David K.  
 APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 FILE REFERENCE: Acemica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761  
 CURRENT FILING DATE: 2001-05-23  
 PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 37202  
 ; LENGTH: 43  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AL035530.11  
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.86  
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1  
 ; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.7  
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5  
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.95  
 ; OTHER INFORMATION: SWISSPROT HIT: P33485, EVALUE 3.60e+00  
 ; OTHER INFORMATION: EST\_HUMAN HIT: A1765888.1, EVALUE 3.00e-18  
 ; US-09-864-761-37202

Query Match 95.0%; Score 19; DB 9; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 ||||  
 Db 40 VAEF 43

## RESULT 22

US-10-424-599-183032  
 Sequence 183032, Application US/10424599  
 Publication No. US20040031072A1

## GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J  
 APPLICANT: Kovalic David K  
 APPLICANT: Zhou Yihua  
 APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 183032  
LENGTH: 43  
TYPE: PRT  
ORGANISM: Glycine max

FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_136291C.1.pap  
-10-424-599-183032

Query Match 95.0%; Score 19; DB 12; Length 43;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||||  
25 VAEF 28

SULT 23  
-10-424-599-171318  
Sequence 171318, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 171318  
LENGTH: 45  
TYPE: PRT  
ORGANISM: Glycine max

FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_125715C.1.pap  
-10-424-599-171318

Query Match 95.0%; Score 19; DB 12; Length 45;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||||  
20 VAEF 23

SULT 24  
-09-864-761-38166  
Sequence 38166, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wenheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6

;; PRIOR FILING DATE: 2000-10-04  
;; PRIOR APPLICATION NUMBER: US 60/236,359  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/00666  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00667  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 38166  
;; LENGTH: 48  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AL049833.3  
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1  
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96  
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1  
;; OTHER INFORMATION: EXPRESSED IN HB1100, SIGNAL = 1.1  
;; OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.2  
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1  
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1  
;; OTHER INFORMATION: SWISSPROT HIT: P11298, EVALU 7.90e+00  
US-09-864-761-38166

Query Match 95.0%; Score 19; DB 9; Length 48;  
Best Local Similarity 100.0%; Pred. No. 6.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
|||||  
Db 3 VAEF 6

RESULT 25  
US-10-424-599-205632  
Sequence 205632, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 205632  
LENGTH: 52

TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_27713C.1.pap  
S-10-424-599-205632

Query Match 95.0%; Score 19; DB 12; Length 52;  
Best Local Similarity 100.0%; Pred. No. 6.9e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
b 26 VAEF 29

RESULT 26  
S-10-424-599-192247  
Sequence 192247, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 192247  
LENGTH: 55  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_1561C.1.pap  
S-10-424-599-192247

Query Match 95.0%; Score 19; DB 12; Length 55;  
Best Local Similarity 100.0%; Pred. No. 7.3e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
b 50 VAEF 53

RESULT 27  
S-10-424-599-233502  
Sequence 233502, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 233502  
LENGTH: 55  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_52879C.1.pap  
S-10-424-599-233502

Query Match 95.0%; Score 19; DB 12; Length 55;  
Best Local Similarity 100.0%; Pred. No. 7.3e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 24 VAEF 27

RESULT 28  
US-10-372-003A-63  
Sequence 63, Application US/10372003A  
Publication No. US20030215846A1  
GENERAL INFORMATION:  
APPLICANT: Watt, Paul  
APPLICANT: Thomas, Richard  
APPLICANT: Hopkins, Richard  
TITLE OF INVENTION: Methods of constructing and screening  
TITLE OF INVENTION: diverse expression libraries  
FILE REFERENCE: FBIC40.001CP1  
CURRENT APPLICATION NUMBER: US/10/372,003A  
CURRENT FILING DATE: 2003-02-21  
PRIOR APPLICATION NUMBER: US 09/568,229  
PRIOR FILING DATE: 2000-05-05  
PRIOR APPLICATION NUMBER: US 60/132,711  
PRIOR FILING DATE: 1999-05-05  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 63  
LENGTH: 55  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic clone BGF05 with first open reading  
OTHER INFORMATION: frame  
US-10-372-003A-63

Query Match 95.0%; Score 19; DB 15; Length 55;  
Best Local Similarity 100.0%; Pred. No. 7.3e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 28 VAEF 31

RESULT 29  
US-10-372-003A-67  
Sequence 67, Application US/10372003A  
Publication No. US20030215846A1  
GENERAL INFORMATION:  
APPLICANT: Watt, Paul  
APPLICANT: Thomas, Wayne  
APPLICANT: Hopkins, Richard  
TITLE OF INVENTION: Methods of constructing and screening  
TITLE OF INVENTION: diverse expression libraries  
FILE REFERENCE: FBIC40.001CP1  
CURRENT APPLICATION NUMBER: US/10/372,003A  
CURRENT FILING DATE: 2003-02-21  
PRIOR APPLICATION NUMBER: US 09/568,229  
PRIOR FILING DATE: 2000-05-05  
PRIOR APPLICATION NUMBER: US 60/132,711  
PRIOR FILING DATE: 1999-05-05  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 67  
LENGTH: 55  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic clone BGF06 with first open reading  
OTHER INFORMATION: frame  
US-10-372-003A-67

Query Match 95.0%; Score 19; DB 15; Length 55;  
Best Local Similarity 100.0%; Pred. No. 7.3e+02; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||||  
28 VAEF 31

SULT 30  
-10-219-329-4  
Sequence 4, Application US/10219329  
Publication No. US20030096757A1  
GENERAL INFORMATION:  
APPLICANT: Quirk, Stephen  
TITLE OF INVENTION: Anti-Cancer and Wound Healing Compounds  
FILE REFERENCE: 1443.035W01  
CURRENT APPLICATION NUMBER: US/10/219,329  
CURRENT FILING DATE: 2002-08-15  
PRIOR APPLICATION NUMBER: US 10/032,376  
PRIOR FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: US 60/312,726  
PRIOR FILING DATE: 2001-08-16  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 56  
TYPE: PRT  
ORGANISM: Homo sapiens  
-10-219-329-4

Query Match 95.0%; Score 19; DB 14; Length 56;  
Best Local Similarity 100.0%; Pred. No. 7.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||||  
31 VAEF 34

SULT 31  
-10-153-185-4  
Sequence 4, Application US/10153185  
Publication No. US20030148959A1  
GENERAL INFORMATION:  
APPLICANT: Quirk, Stephen  
TITLE OF INVENTION: Anti-Aging and Wound Healing Compounds  
FILE REFERENCE: 1443.034U01  
CURRENT APPLICATION NUMBER: US/10/153,185  
CURRENT FILING DATE: 2002-08-13  
PRIOR APPLICATION NUMBER: US 10/032,376  
PRIOR FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: US 60/312,726  
PRIOR FILING DATE: 2001-08-16  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 56  
TYPE: PRT  
ORGANISM: Homo sapiens  
-10-153-185-4

Query Match 95.0%; Score 19; DB 14; Length 56;  
Best Local Similarity 100.0%; Pred. No. 7.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
|||||  
31 VAEF 34

SULT 32  
-10-219-561-4

; Sequence 4, Application US/10219561  
; Publication No. US20030166567A1  
; GENERAL INFORMATION:  
; APPLICANT: Quirk, Stephen  
; APPLICANT: Malik, Sohail  
; APPLICANT: Villanueva, Julie M.  
; TITLE OF INVENTION: Anti-Aging and Wound Healing Compounds  
; FILE REFERENCE: 1443.008US2  
; CURRENT APPLICATION NUMBER: US/10/219,561  
; CURRENT FILING DATE: 2002-08-15  
; PRIOR APPLICATION NUMBER: US 10/032,376  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: US 10/153,185  
; PRIOR FILING DATE: 2002-05-21  
; PRIOR APPLICATION NUMBER: US 60/312,726  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 56  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-219-561-4

Query Match 95.0%; Score 19; DB 14; Length 56;  
Best Local Similarity 100.0%; Pred. No. 7.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
|||||  
Db 31 VAEF 34

RESULT 33  
US-10-424-599-215734  
; Sequence 215734, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 215734  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_36836C.1.pap  
US-10-424-599-215734

Query Match 95.0%; Score 19; DB 12; Length 57;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
|||||  
Db 48 VAEF 51

RESULT 34  
US-10-424-599-260762  
; Sequence 260762, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua



APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 260762  
LENGTH: 57  
TYPE: PRT  
ORGANISM: Glycine max

FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_27490C.1.pep  
S-10-424-599-260762

Query Match 95.0%; Score 19; DB 12; Length 57;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
b 4 VAEF 7

## RESULT 35

S-10-372-003A-75  
Sequence 75, Application US/10372003A  
Publication No. US20030215846A1

## GENERAL INFORMATION:

APPLICANT: Watt, Paul

APPLICANT: Thomas, Wayne

APPLICANT: Hopkins, Richard

TITLE OF INVENTION: Methods of constructing and screening

FILE REFERENCE: diverse expression libraries

CURRENT APPLICATION NUMBER: US/10/372,003A

CURRENT FILING DATE: 2003-02-21

PRIOR APPLICATION NUMBER: US 09/568,229

PRIOR FILING DATE: 2000-05-05

PRIOR APPLICATION NUMBER: US 60/132,711

PRIOR FILING DATE: 1999-05-05

NUMBER OF SEQ ID NOS: 81

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 75

LENGTH: 57

TYPE: PRT

ORGANISM: Artificial Sequence

## FEATURE:

OTHER INFORMATION: synthetic clone BGF24 with first open reading

OTHER INFORMATION: frame

S-10-372-003A-75

Query Match 95.0%; Score 19; DB 15; Length 57;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
b 28 VAEF 31

## RESULT 36

S-10-424-599-214930

Sequence 214930, Application US/10424599

Publication No. US20040031072A1

## GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J

APPLICANT: Zhou Yihua

APPLICANT: Kovalic David K

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 214930  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT3847\_36107C.1.pep  
US-10-424-599-214930

Query Match 95.0%; Score 19; DB 12; Length 58;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
Db 13 VAEF 16

## RESULT 37

US-09-864-761-45291

Sequence 45291, Application US/09864761

Patent No. US20020048763A1

## GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aemica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 45291

LENGTH: 60  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC009079.4  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.54  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.77  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.65  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.74  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.48  
OTHER INFORMATION: EST HUMAN HIT: AV756022.1, EVALUE 6.00e-28  
OTHER INFORMATION: SWISSPROT HIT: P56315, EVALUE 2.00e+00  
-09-864-761-45291

Query Match 95.0%; Score 19; DB 9; Length 60;  
Best Local Similarity 100.0%; Pred. No. 8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 VAEF 5  
8 VAEF 11

SULT 38  
-09-393-634-80  
Sequence 80, Application US/09393634  
Patent No. US20020051997A1  
GENERAL INFORMATION:  
APPLICANT: Zuker, Charles S.  
APPLICANT: Adler, Jon Elliot  
APPLICANT: Ryba, Nick  
APPLICANT: Mueller, Ken  
APPLICANT: Hoon, Mark  
APPLICANT: The Regents of the University of California  
APPLICANT: The Government of the United States of America  
APPLICANT: as represented by the Secretary of the  
APPLICANT: Department of Health and Human Services  
TITLE OF INVENTION: SF, a No. US20020051997A1el Family of Taste Receptors  
FILE REFERENCE: 02307E-098000US  
CURRENT APPLICATION NUMBER: US/09/393,634  
CURRENT FILING DATE: 1999-09-10  
NUMBER OF SEQ ID NOS: 92  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 80  
LENGTH: 68  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: human GR24  
-09-393-634-80

Query Match 95.0%; Score 19; DB 9; Length 68;  
Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 VAEF 5  
14 VAEF 17

SULT 39  
-09-864-408A-40  
Sequence 40, Application US/09864408A  
Publication No. US2004000947A1  
GENERAL INFORMATION:  
APPLICANT: Leach, Martin D.  
APPLICANT: Shmukets, Richard A.  
TITLE OF INVENTION: No. US2004000947A1el Human Polynucleotides and Polypeptides Encod  
FILE REFERENCE: 21402-012  
CURRENT APPLICATION NUMBER: US/09/864,408A  
CURRENT FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 60/206,690  
PRIOR FILING DATE: 2000-05-24

NUMBER OF SEQ ID NOS: 9068  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 40  
LENGTH: 68  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-864-408A-40  
Query Match 95.0%; Score 19; DB 11; Length 68;  
Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 3 VAEF 6

RESULT 40  
US-10-364-861-80  
Sequence 80, Application US/10364861  
Publication No. US20040038312A1  
GENERAL INFORMATION:  
APPLICANT: Zuker, Charles S.  
APPLICANT: Adler, Jon Elliot  
APPLICANT: Hoon, Mark  
APPLICANT: Ryba, Nick  
APPLICANT: Mueller, Ken  
APPLICANT: The Regents of the University of California  
APPLICANT: The Government of the United States of America  
APPLICANT: as represented by the Secretary of the  
APPLICANT: Department of Health and Human Services  
TITLE OF INVENTION: T2R, a No. US20040038312A1el Family of Taste Receptors  
FILE REFERENCE: 02307E-098020US  
CURRENT APPLICATION NUMBER: US/10/364,861  
CURRENT FILING DATE: 2003-06-30  
PRIOR APPLICATION NUMBER: US 09/393,634  
PRIOR FILING DATE: 1999-09-10  
NUMBER OF SEQ ID NOS: 95  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 80  
LENGTH: 68  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: human T2R24, GR24 or SF24  
US-10-364-861-80

Query Match 95.0%; Score 19; DB 12; Length 68;  
Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 VAEF 5  
14 VAEF 17

RESULT 41  
US-10-383-982-80  
Sequence 80, Application US/10383982  
Publication No. US20030157568A1  
GENERAL INFORMATION:  
APPLICANT: Zuker, Charles S.  
APPLICANT: Adler, Jon Elliot  
APPLICANT: Ryba, Nick  
APPLICANT: Mueller, Ken  
APPLICANT: Hoon, Mark  
APPLICANT: The Regents of the University of California  
APPLICANT: The Government of the United States of America  
APPLICANT: as represented by the Secretary of the  
APPLICANT: Department of Health and Human Services  
TITLE OF INVENTION: SF, a No. US20030157568A1el Family of Taste Receptors  
FILE REFERENCE: 02307E-098000US  
CURRENT APPLICATION NUMBER: US/10/383,982

CURRENT FILING DATE: 2003-03-07  
PRIOR APPLICATION NUMBER: US/09/393,634  
PRIOR FILING DATE: 1999-09-10  
NUMBER OF SEQ ID NOS: 92  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 80  
LENGTH: 68  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: human GR24  
S-10-383-982-80  
Query Match 95.0%; Score 19; DB 14; Length 68;  
Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Y 2 VAEF 5  
b 14 VAEF 17  
|||||  
RESULT 42  
S-09-864-761-45716  
Sequence 45716, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aecmics-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 45716  
LENGTH: 69  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL162171.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.81  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.91  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.85  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.72  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.71  
OTHER INFORMATION: EST HUMAN HIT: BE559987.1, EVALUATE 2.00e-03  
OTHER INFORMATION: SWISSPROT HIT: P56093, EVALUATE 2.00e+00  
US-09-864-761-45716  
Query Match 95.0%; Score 19; DB 9; Length 69;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 15 VAEF 18  
|||||  
RESULT 43  
US-09-738-626-5596  
Sequence 5596, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAOKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 5596  
LENGTH: 69  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-5596  
Query Match 95.0%; Score 19; DB 9; Length 69;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 23 VAEF 26  
|||||  
RESULT 44  
US-10-424-599-185385  
Sequence 185385, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 185385  
LENGTH: 69  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_138417C.1.pap  
US-10-424-599-185385

Query Match 95.0%; Score 19; DB 12; Length 69;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 24 VAEF 27

RESULT 45  
US-10-029-386-28645  
Sequence 28645, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
SEQ ID NO 28645  
LENGTH: 70  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO CHR17.1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.8  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.3  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.8  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.7  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.7  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8.7  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7  
OTHER INFORMATION: SWISSPROT HIT: P78716, EVALUATE 3.00e-04  
US-10-029-386-28645

Query Match 95.0%; Score 19; DB 14; Length 70;  
Best Local Similarity 100.0%; Pred. No. 9.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 43 VAEF 46

RESULT 46  
US-10-424-599-202605  
Sequence 202605, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 202605  
LENGTH: 71  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_24978C.1.pap  
US-10-424-599-202605

Query Match 95.0%; Score 19; DB 12; Length 71;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 17 VAEF 20

RESULT 47  
US-10-424-599-274981  
Sequence 274981, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 274981  
LENGTH: 71  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(71)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_90329C.1.pap  
US-10-424-599-274981

Query Match 95.0%; Score 19; DB 12; Length 71;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 39 VAEF 42

RESULT 48  
US-10-029-386-30348  
Sequence 30348, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
FILE REFERENCE: AEOMICA-X-2

```

CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 30348
LENGTH: 71
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR7.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.82
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
OTHER INFORMATION: SWISSPROT HIT: P21333, EVALUATE 5.00e-22
S-10-029-386-30348

Query Match          95.0%; Score 19; DB 14; Length 71;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y      2 VAEF 5
      |||||
b      60 VAEF 63

RESULT 49
S-10-029-386-29320
Sequence 29320, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David X.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 29320
LENGTH: 72
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR2.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.98
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.99
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.6
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.81
OTHER INFORMATION: SWISSPROT HIT: P14650, EVALUATE 6.00e-14
S-10-029-386-29320

Query Match          95.0%; Score 19; DB 14; Length 72;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y      2 VAEF 5
      |||||
b      45 VAEF 48

RESULT 50
S-10-214-188-10
Sequence 10, Application US/10214188
Publication No. US2003002260A1
GENERAL INFORMATION:
APPLICANT: LA THANGUE, NICHOLAS B.

/
/ BERNARDS, RENE
/ HIJLMANS, ELEANORE M.
/ TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: NIXON & VANDERHYE P.C.
/ STREET: 1100 NORTH GLEBE ROAD
/ CITY: ARLINGTON
/ STATE: VIRGINIA
/ COUNTRY: U.S.A.
/ ZIP: 22201-4714
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/214,188
/ FILING DATE: 08-Aug-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/894,139
/ FILING DATE: 13-AUG-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: WILSON, MARY J.
/ REGISTRATION NUMBER: 32,955
/ REFERENCE/DOCKET NUMBER: 620-22
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 816-4000
/ TELEFAX: (703) 816-4100
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 74 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ SEQUENCE DESCRIPTION: SEQ ID NO: 10:
/ US-10-214-188-10
/
/ Query Match          95.0%; Score 19; DB 14; Length 74;
/ Best Local Similarity 100.0%; Pred. No. 9.9e+02;
/ Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
/ QY      2 VAEF 5
/      |||||
/ Db      23 VAEF 26
/
/ RESULT 51
/ US-10-156-761-9071
/ Sequence 9071, Application US/10156761
/ Publication No. US20030119018A1
/ GENERAL INFORMATION:
/ APPLICANT: OMURA, SATOSHI
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: ISHIKAWA, JUN
/ APPLICANT: HORIKAWA, HIROSHI
/ APPLICANT: SHIBA, TADAYOSHI
/ APPLICANT: SAKAKI, YOSHIYUKI
/ APPLICANT: HATTORI, MASAHIRA
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-262
/ CURRENT APPLICATION NUMBER: US/10/156,761
/ CURRENT FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ NUMBER OF SEQ ID NOS: 15109
/ SEQ ID NO 9071
/ LENGTH: 74
/ TYPE: PRT
/
```

ORGANISM: Streptomyces avermitilis  
-10-156-761-9071

Query Match 95.0%; Score 19; DB 14; Length 74;  
Best Local Similarity 100.0%; Pred. No. 9.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 36 VAEF 39

## SULT 52

-10-424-599-156128  
Sequence 156128, Application US/10424599  
Publication No. US20040031072A1

## GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 156128  
LENGTH: 75  
TYPE: PRT  
ORGANISM: Glycine max  
NAME/KEY: unsure  
LOCATION: (1)..(75)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_112004C.1.pep

## -10-424-599-156128

Query Match 95.0%; Score 19; DB 12; Length 75;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 2 VAEF 5

## SULT 53

-10-424-599-261424  
Sequence 261424, Application US/10424599  
Publication No. US20040031072A1

## GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 261424  
LENGTH: 75  
TYPE: PRT  
ORGANISM: Glycine max  
NAME/KEY: unsure  
LOCATION: (1)..(75)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_78089C.1.pep

## -10-424-599-261424

Query Match 95.0%; Score 19; DB 12; Length 75;  
Best Local Similarity 100.0%; Pred. No. 1e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 12 VAEF 15

## RESULT 54

US-10-424-599-253204  
Sequence 253204, Application US/10424599  
Publication No. US20040031072A1

## GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 253204  
LENGTH: 76  
TYPE: PRT  
ORGANISM: Glycine max  
NAME/KEY: unsure  
LOCATION: (1)..(76)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_70669C.1.pep

## US-10-424-599-253204

Query Match 95.0%; Score 19; DB 12; Length 76;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 64 VAEF 67

## RESULT 55

US-10-424-599-211908  
Sequence 211908, Application US/10424599  
Publication No. US20040031072A1

## GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 211908  
LENGTH: 77  
TYPE: PRT  
ORGANISM: Glycine max  
NAME/KEY: unsure  
LOCATION: (1)..(77)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_3337C.1.pep

## US-10-424-599-211908

Query Match 95.0%; Score 19; DB 12; Length 77;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
Db 16 VAEF 19

RESULT 56  
S-10-335-977-7493  
Sequence 7493, Application US/10335977  
Publication No. US20040052799A1  
GENERAL INFORMATION:  
APPLICANT: DOUGLAS SMITH et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
RELATING TO HELICOBACTER PYLORI FOR  
DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 10031  
CORRESPONDENCE ADDRESS:  
ADDRESSER: LAHIVE & COCKFIELD  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: Windows NT 4.0  
SOFTWARE: UNIX  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/335,977  
FILING DATE: 30-Dec-2002  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/993,002  
FILING DATE: 17-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: GTN-018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 7493:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 77 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Helicobacter pylori  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...77  
SEQUENCE DESCRIPTION: SEQ ID NO: 7493:  
S-10-335-977-7493  
Query Match 95.0%; Score 19; DB 12; Length 77;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Y 2 VAEF 5  
O 40 VAEF 43

RESULT 57  
S-10-029-386-32709  
Sequence 32709, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 32709  
; LENGTH: 77  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC004186.1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.76  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9  
; OTHER INFORMATION: SWISSPROT HIT: P14373, EVALUE 7.50e+00  
US-10-029-386-32709  
Query Match 95.0%; Score 19; DB 14; Length 77;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 11 VAEF 14

RESULT 58  
US-10-424-599-162039  
; Sequence 162039, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 162039  
; LENGTH: 78  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_117339C.1.pap  
US-10-424-599-162039  
Query Match 95.0%; Score 19; DB 12; Length 78;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 28 VAEF 31

RESULT 59  
US-10-424-599-262701  
; Sequence 262701, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 262701  
; LENGTH: 78

```
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_79240C.1.pap
-10-424-599-262701

Query Match          95.0%; Score 19; DB 12; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      2 VAEF 5
      ||||
      5 VAEF 8

-10-363-616-368
Sequence 368, Application US/10363616
Publication No. US20040044181A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-113 (793)
CURRENT APPLICATION NUMBER: US/10/363,616
CURRENT FILING DATE: 2003-03-03
PRIOR APPLICATION NUMBER: 09/654,935
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 490
SEQ ID NO 368
LENGTH: 79
TYPE: PRT
ORGANISM: Homo sapiens
-10-363-616-368

Query Match          95.0%; Score 19; DB 12; Length 79;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      2 VAEF 5
      ||||
      24 VAEF 27

-10-424-599-216097
Sequence 216097, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 216097
LENGTH: 80
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_37165C.1.pap
-10-424-599-216097

Query Match          95.0%; Score 19; DB 12; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      2 VAEF 5
      ||||
      35 VAEF 38
```

```
RESULT 62
US-10-367-980A-14
; Sequence 14, Application US/10367980A
; Publication No. US20030228592A1
; GENERAL INFORMATION:
; APPLICANT: St Vincent's Institute of Medical Research
; APPLICANT: Rogers, Suzanne D
; APPLICANT: Best, James D
; TITLE OF INVENTION: Human Facilitative Glucose Transport Protein GLUT8
; FILE REFERENCE: VS:AJH.FP47928
; CURRENT APPLICATION NUMBER: US/10/367,980A
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 09/509,731
; PRIOR FILING DATE: 2003-02-19
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: GLUT2
US-10-367-980A-14

Query Match          95.0%; Score 19; DB 15; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
DB      ||||
      13 VAEF 16

RESULT 63
US-09-864-408A-5282
; Sequence 5282, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Encoc
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5282
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-5282

Query Match          95.0%; Score 19; DB 11; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VAEF 5
DB      ||||
      52 VAEF 55

RESULT 64
US-10-078-090-125
; Sequence 125, Application US/10078090
; Publication No. US20030044815A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
```



```

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 201853
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(85)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_24299C.1.pap
US-10-424-599-201853

Query Match          95.0%; Score 19; DB 12; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 VAEF 5
Db      58 VAEF 61

RESULT 67
US-09-815-242-11224
; Sequence 11224, Application US/09815242
; Patent No. US20020081569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlseen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11224
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11224

Query Match          95.0%; Score 19; DB 9; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 VAEF 5
Db      12 VAEF 15

```

SULT 68

-09-815-242-11245  
 Sequence 11245, Application US/09815242  
 Patent No. US20020061569A1  
 GENERAL INFORMATION:  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari L.  
 APPLICANT: Zyskind, Judith W.  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John D.  
 APPLICANT: Carr, Grant J.  
 APPLICANT: Yamamoto, Robert T.  
 APPLICANT: Xu, H. Howard  
 TITLE OF INVENTION: Identification of Essential Genes in  
 FILE REFERENCE: ELITRA 011A  
 CURRENT APPLICATION NUMBER: US/09/815,242  
 PRIOR FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 NUMBER OF SEQ ID NOS: 14110  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 11245  
 LENGTH: 89  
 TYPE: PRT  
 ORGANISM: Haemophilus influenzae

Query Match 95.0%; Score 19; DB 9; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 VAEF 5  
 ||||  
 12 VAEF 15

SULT 69

-10-282-122A-58489  
 Sequence 58489, Application US/10282122A  
 Publication No. US20040029129A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari  
 APPLICANT: Zyskind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: ELITRA 034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636

;; PRIOR FILING DATE: 2000-05-23  
 ;; PRIOR APPLICATION NUMBER: 60/207,727  
 ;; PRIOR FILING DATE: 2000-05-26  
 ;; PRIOR APPLICATION NUMBER: 60/230,335  
 ;; PRIOR FILING DATE: 2000-09-06  
 ;; PRIOR APPLICATION NUMBER: 60/230,347  
 ;; PRIOR FILING DATE: 2000-09-09  
 ;; PRIOR APPLICATION NUMBER: 60/242,578  
 ;; PRIOR FILING DATE: 2000-10-23  
 ;; PRIOR APPLICATION NUMBER: 60/253,625  
 ;; PRIOR FILING DATE: 2000-11-27  
 ;; PRIOR APPLICATION NUMBER: 60/257,931  
 ;; PRIOR FILING DATE: 2000-12-22  
 ;; PRIOR APPLICATION NUMBER: 60/267,636  
 ;; PRIOR FILING DATE: 2001-02-09  
 ;; PRIOR APPLICATION NUMBER: 60/269,308  
 ;; PRIOR FILING DATE: 2001-02-16  
 ;; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ;; NUMBER OF SEQ ID NOS: 78614  
 ;; SOFTWARE: PatentIn version 3.1  
 ;; SEQ ID NO 58489  
 ;; LENGTH: 89  
 ;; TYPE: PRT  
 ;; ORGANISM: Haemophilus influenzae

US-10-282-122A-58489

Query Match 95.0%; Score 19; DB 12; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VAEF 5  
 Db 12 VAEF 15

RESULT 70

US-10-282-122A-58530  
 Sequence 58530, Application US/10282122A  
 Publication No. US20040029129A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari  
 APPLICANT: Zyskind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: ELITRA 034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 58530  
LENGTH: 89

TYPE: PRT  
ORGANISM: Haemophilus influenzae  
S-10-282-122A-58530

Query Match 95.0%; Score 19; DB 12; Length 89;  
Best Local Similarity 100.0%; Pred.No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 VAEF 5  
12 VAEF 15

## RESULT 71

S-10-282-122A-66960  
Sequence 66960, Application US/10282122A  
Publication No. US20040029129A1

## GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.

## TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 66960

LENGTH: 89

TYPE: PRT

ORGANISM: Pasteurella multocida

S-10-282-122A-66960

Query Match 95.0%; Score 19; DB 12; Length 89;  
Best Local Similarity 100.0%; Pred.No. 1.2e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VAEF 5  
DB 12 VAEF 15

## RESULT 72

US-10-282-122A-68774  
Sequence 68774, Application US/10282122A  
Publication No. US20040029129A1

## GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.

## TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 68774

LENGTH: 89

TYPE: PRT

ORGANISM: Proteus mirabilis

US-10-282-122A-68774

Query Match 95.0%; Score 19; DB 12; Length 89;  
Best Local Similarity 100.0%; Pred.No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAEF 5  
DB 12 VAEF 15

## RESULT 73

US-10-424-599-180307

Sequence 180307, Application US/10424599

Publication No. US20040031072A1

## GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 180307  
LENGTH: 89  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_133832C.1.pep  
-10-424-599-180307

Query Match 95.0%; Score 19; DB 12; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 VAEF 5  
11 VAEF 14

RESULT 74  
-10-097-111-360  
Sequence 360, Application US/10097111  
Publication No. US20030138771A1  
GENERAL INFORMATION:  
APPLICANT: PELLETIER, JERRY  
APPLICANT: GROS, PHILIPPE  
APPLICANT: DUBOW, MICHAEL  
TITLE OF INVENTION: DNA SEQUENCES FROM S. PNEUMONIAE BACTERIOPHAGE BP1 THAT  
TITLE OF INVENTION: ENCODE ANTI-MICROBIAL POLYPEPTIDES  
FILE REFERENCE: 073406-0603  
CURRENT APPLICATION NUMBER: US/10/097,111  
CURRENT FILING DATE: 2002-07-24  
PRIOR APPLICATION NUMBER: 09/676,412  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/157,218  
PRIOR FILING DATE: 1999-09-30  
NUMBER OF SEQ ID NOS: 552  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 360  
LENGTH: 89  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
-10-097-111-360

Query Match 95.0%; Score 19; DB 14; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 VAEF 5  
11 VAEF 14

RESULT 75  
-10-424-599-245830  
Sequence 245830, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 245830  
LENGTH: 90  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_64015C.1.pep  
US-10-424-599-245830

Query Match 95.0%; Score 19; DB 12; Length 90;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 VAEF 5  
60 VAEF 63

Search completed: May 24, 2004, 17:42:32  
Job time : 55.6429 secs